

ITALIAN PHYSICAL SOCIETY

PROCEEDINGS
OF THE
INTERNATIONAL SCHOOL OF PHYSICS
« ENRICO FERMI »

COURSE XIX

edited by B. PETERS
Director of the Course

VARENNA ON LAKE COMO
VILLA MONASTERO

MAY 23 - JUNE 3 1961

Cosmic rays, solar particles and space research

1963



ACADEMIC PRESS • NEW YORK AND LONDON

SOCIETÀ ITALIANA DI FISICA

RENDICONTI
DELLA
SCUOLA INTERNAZIONALE DI FISICA
« ENRICO FERMI »

XIX CORSO

a cura di B. PETERS
Direttore del Corso

VARENNA SUL LAGO DI COMO
VILLA MONASTERO

23 MAGGIO - 3 GIUGNO 1961

Raggi cosmici, particelle solari e ricerca spaziale

1963



ACADEMIC PRESS • NEW YORK AND LONDON

IIA LIB.

ACADEMIC PRESS INC.
111 FIFTH AVENUE
NEW YORK 3, N. Y.

United Kingdom Edition
Published by
ACADEMIC PRESS INC. (LONDON) LTD.
BERKELEY SQUARE HOUSE, LONDON W. 1

COPYRIGHT © 1963, BY SOCIETÀ ITALIANA DI FISICA

ALL RIGHTS RESERVED

NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM,
BY PHOTOSTAT, MICROFILM, OR ANY OTHER MEANS,
WITHOUT WRITTEN PERMISSION FROM THE PUBLISHERS.

Library of Congress Catalog Card Number: 63-14743

PRINTED IN ITALY

INDICE

B. PETERS - Prolusione al Corso	pag. IX
Gruppo fotografico dei partecipanti al Corso	fuori testo

Relazioni tra Sole e Terra.

M. HACK - The surface of the sun	pag. 1
J. H. REID - An optical phenomenon associated with type IV solar radio emission	» 14
H. ELLIOT - Motion of cosmic ray particles in the geomagnetic field .	» 16
K. W. OGILVIE - Solar protons	» 28
C. S. WARWICK - Propagation of solar particles and the interstellar magnetic field	» 61
E. R. HARRISON - Solar corpuscular radiation and the nature of the earth's immediate interplanetary environment	» 65
E.-Å. BRUNBERG - The production of high-energy particles in the sun and in the solar system	» 85
C.-G. FÄLTHAMMAR - Acceleration of cosmic-ray particles by magnetic pumping	» 99
E.-Å. BRUNBERG - The earth's radiation belts and the aurora	» 113
R. E. GENDRIN - Mechanisms of V.L.F. emission	» 125

Problemi galattici.

C. J. WADDINGTON - The primary cosmic radiation	» 135
J. LABEYRIE - Use of solid-state detectors as energy spectrometers for cosmic high-energy particles	» 187
L. SCARSI - Extensive air showers	» 194
B. DAYTON - Electrons in the galactic cosmic radiation	» 200
J. GOUPIL - γ -spectrography in a rocket	» 211
H. C. VAN DE HULST - The interstellar gas and the galactic halo . .	» 220
J. M. GREENBERG - The polarization of starlight in the interstellar medium	» 227
J. M. GREENBERG - Optical astronomy above the earth's atmosphere .	» 235

Storia ed origine dei raggi cosmici.

J. GEISS, H. OESCHGER and U. SCHWARZ – The history of cosmic radiation as revealed by isotopic changes in the meteorites and on the earth	pag. 247
J. ZÄHRINGER – Some recent investigations on primordial rare gases and ^{129}Xe in meteorites	» 275
F. HOYLE – On the origin of cosmic rays	» 293

Strumenti e programmi di ricerca.

G. H. LUDWIG – The U. S. program for particles and fields measurements in space	» 320
G. H. LUDWIG – Spacecraft information systems	» 348
E. B. DORLING – The British Space Research Programme	» 386
H. ELLIOT – Cosmic ray measurements in the U.S./U.K. Satellite S-51	» 392
E.-Å. BRUNBERG – Research and research plans at Kiruna geophysical observatory	» 402
M. CECCARELLI – Programs of radioastronomy in Italy	» 408
H. C. VAN DE HULST – International agreement as a vital factor in radio astronomy and space research	» 412

Prolusione al Corso.

B. PETERS

Direttore del Corso

It is a pleasure to thank the Mayor of Varenna for the hospitality extended to us. I would like to thank Professor POLVANI for his words of welcome and through him thank the Italian Physical Society for its generosity which makes it possible to hold Summer Courses in this most beautiful place.

I welcome all of you who have come to participate in this course and shall use this opportunity to say a few words about its origin and purpose.

Last Fall, here in the Villa Monastero, Professor POLVANI, Prof. HOUTERMANS and I discussed the possibility and usefulness of organizing a course on Cosmic Radiation, Solar Particles and Space Research. Let me briefly explain the reason for choosing this subject.

As you know, Cosmic Ray Physics became an extremely active field of research soon after the war. Earlier, it had been a specialized field of study, almost completely isolated from other sciences; but in the last 12 years it has branched out.

Through the discovery and the study of the complex nature of the primary cosmic radiation it has become an integral part of astrophysics.

Through the study of cosmic ray produced isotopes it has made valuable contributions to fields of science as far apart as oceanography, meteorology and archaeology, and has contributed to our knowledge of the origin of meteorites.

The analysis of collisions of cosmic ray particles in the atmosphere has made it possible for the first time to study nuclear and electromagnetic interactions at energies which are large compared to the rest mass of the particles involved. Cosmic ray physics, thereby, has given rise to a new and still very active branch of physics, the physics of elementary particles.

Cosmic ray physicists are responsible for the discovery of mesons and of almost all the so-called strange particles. The results which they have obtained led to the design of the powerful particle accelerators which are in operation or in the course of construction in many parts of the world.

Perhaps the most active period of cosmic ray physics and the period which produced the most far-reaching changes in our views on the properties of matter

was the period between 1951 and 1953, when new mesons and hyperons were identified and when the processes of their production and decay were studied for the first time.

After 1953 large accelerators and bubble chambers started to operate; there began a less productive period for cosmic ray physics. The study of elementary particles and the study of many electromagnetic and nuclear phenomena at high energies reverted to the laboratory.

There remained many unsolved problems in the nature of the primary cosmic radiation, but they were difficult to handle with the available experimental and technical equipment, consisting mainly of stratosphere balloons, emulsions, cloud chambers, and counters. It became increasingly difficult to improve the resolution between the particles of various mass numbers which compose the primary radiation, to determine the abundance of the rare elements, and to obtain higher statistical accuracy in the energy spectrum of various components.

Also the study of time variations in cosmic ray intensity, which has become of great importance, progressed only slowly during that period, because one had to rely on large networks of co-operating ground stations or on the synchronization of balloon ascents at widely separated places. The most valuable data had to be obtained in the least accessible places, (*i.e.*, the stratosphere in the polar regions) so that every new experiment involved a major expedition.

Recently, during the last three or four years, the rate of progress in cosmic ray physics has again accelerated. New avenues of research have been opened by the techniques of space research. New discoveries on the corpuscular radiation from the galaxy and the sun have been made and more important ones are in the offing. They will be made in the U.S.A. and in the U.S.S.R. but, for the time being not here in Europe, since European cosmic ray physicists have neither satellites nor space probes. If they do not want to continue working with outdated techniques, they have only the choice between switching their field of research and emigration.

About two years ago, Professor AMALDI wrote a memorandum on Europe and space research, which circulated mainly among scientists connected with CERN. In this memorandum Professor AMALDI emphasized that the investigation of space is giving rise to new and important technological developments and that, if Europe made no effort to participate in space research, it may within a period of 20 or 30 years lose its high technological standing and become one of the underdeveloped regions of the world.

About a year ago, the first serious steps were undertaken by individual scientists to explore the possibility of developing space research as a co-operative European effort. The brilliant success of CERN has greatly encouraged this approach of close co-operation between various nations. If such an enterprise was to succeed, it seemed important to counteract the drifting away

of European cosmic ray physicists to other research fields and to other countries. A discussion of the present state of the science with particular emphasis on the problems which have become accessible to investigation by the advent of space vehicles seemed to be a useful step in that direction.

These then were the main considerations governing the choice of subject matter for this course. The scientific programme has been arranged accordingly; it consists of a number of review lectures, each one dealing with a different aspect of solar and galactic particles. These lectures take up about half of the available time. The rest will be devoted to discussions and to shorter contributions from an audience which consists almost entirely of people coming from laboratories which are engaged in one or another aspect of cosmic ray research. I invite all of you who can contribute to the subject matter treated in this course to inform me within the next few days.

On the 1st of December 1960, at Meyrin, the site of CERN, an agreement was signed between representatives of 11 nations (Belgium, Denmark, France, Holland, Italy, Norway, Spain, Switzerland, Sweden, United Kingdom and West Germany) to set up a committee to investigate the possibilities of space research in Europe, and to prepare an international treaty for a European Space Research Organization (ESRO). This committee whose name, « COPERS », derives from the words « Commission Préparatoire Européenne pour la Recherche Spatiale », has held many meetings and is approaching the stage where it can make concrete proposals to the participating governments. It seems likely that these efforts will lead to the setting up of European Space Research laboratories and to the recruiting of staff for these laboratories in the early part of 1962. In anticipation of such an agreement, many countries, now engaged in rocket and upper atmosphere research, have declared their willingness to accept engineers and scientists from other countries for training purposes. Those who are interested in availing themselves of such opportunities should write to Professor P. AUGER, Secretary of COPERS (36 Rue La Pérouse, Paris XVI).

If then it seems likely that Europe will actively carry out space research in the coming years, it becomes important to keep well informed on the development of space research abroad. I am therefore particularly happy that among the participants in our course are two representatives of NASA (National Aeronautic and Space Administration of the U.S.), Dr. LUDWIG and Dr. OGILVIE, who have been active in the most recent work on the properties of cosmic radiation in space.

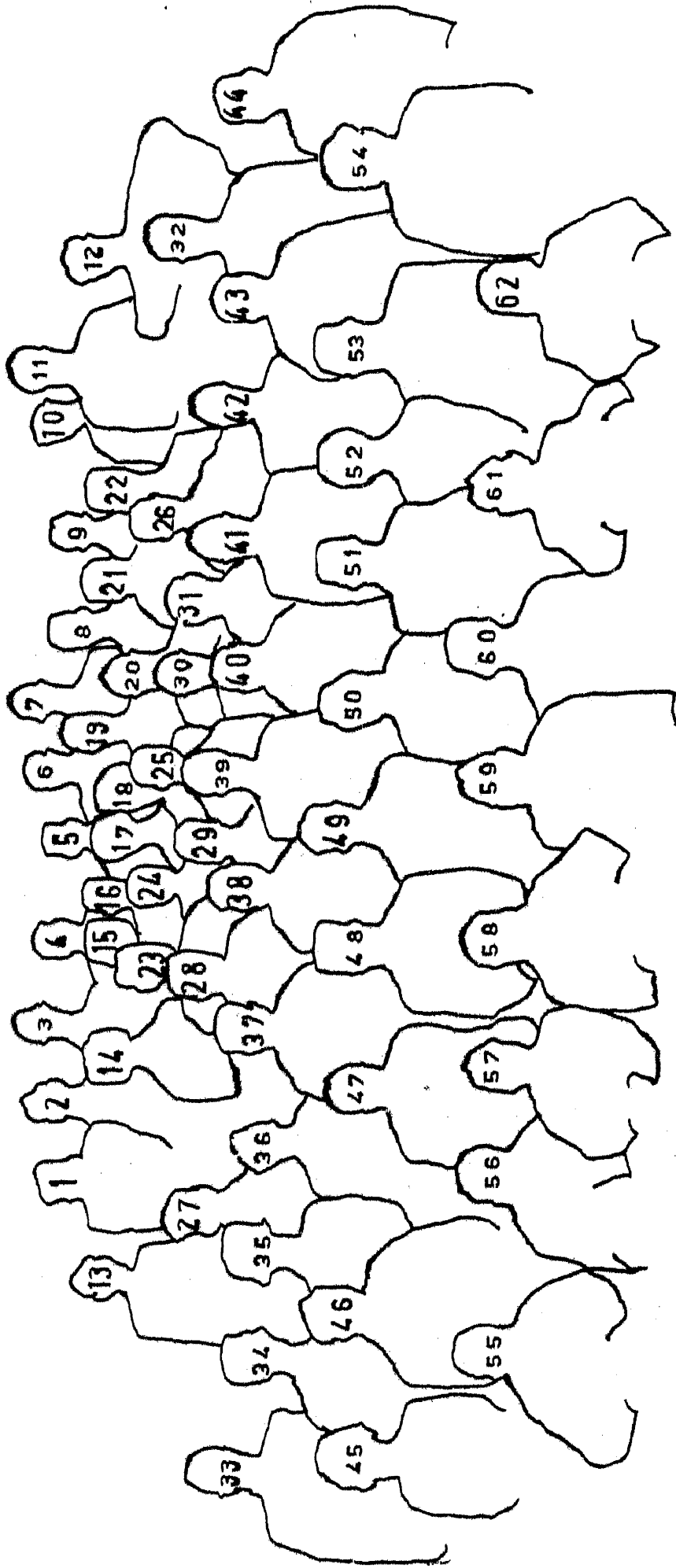
It is a loss to this course that Professor GINZBURG of the Lebedev Institute in Moscow who, as you know, is one of the authors of a detailed and very successful theory on the origin of cosmic radiation, has been unable to accept our invitation.

I also regret to report that Professor DENISSE of the Observatoire de Paris

at Meudon has fallen ill. He had planned to give an introductory lecture the sun from the point of view of an astrophysicist. I am very grateful Professor HACK of the Merate Observatory who, on rather short notice, I agreed to give this introductory lecture.

Before closing let me say a few words about nomenclature. «Cosmic ray» is not a fortunate designation for the mixture of high energy nuclei which originate and are accelerated in our galaxy. «Galactic high energy particle» would be a more appropriate term to distinguish these particles from other forms of radiation which reach us either from our galaxy or from the most distant places of the cosmos. But the name «cosmic rays» has been used for so long that it seems impractical to change it at this late date. This is no reason, however, why we should admit such illogical terms as «solar cosmic radiation» into our vocabulary, in order to designate particles accelerated on or near the sun. This designation has already made its appearance in the literature, but its spread can still be arrested; it should, I feel, be replaced by a less ambiguous term such as «solar corpuscular radiation» or «solar particles» as we have done in the title of this course.

I have the honour of opening now the scientific part of the 1961 program of the «Enrico Fermi» Summer School.



1. J. A. Högbam
2. C. J. Hatton
3. C. G. Falthammar
4. E. Dyring
5. K. Imaeda
6. D. E. Evans
7. J. Goupil
8. M. Tamers
9. T. Matthews
10. H. Elliot
11. N. Iucci
12. E. R. Harrison
13. N. Durgaprasad

14. A. Marzari Chiesa
15. V. Bisi
16. B. Grönkvist
17. G. Ohlsson
18. I. M. Jackerott
19. P. K. F. Grieder
20. K. W. Ogilvie
21. Siitkus
22. K. Goebel
23. G. Tabellini
24. W. Debenedetti
25. A. H. Sautets
26. H. Laster

27. F. Occhionero
28. E. D. Palmatier
29. A. Gandolfi
30. G. R. Stevenson
31. R. J. Sluiz
32. U. Schwarz
33. J. P. Vinti
34. Gheller
35. D. Russo
36. Variola
37. P. Balata
38. G. S. Gokhale
39. C. J. Bland

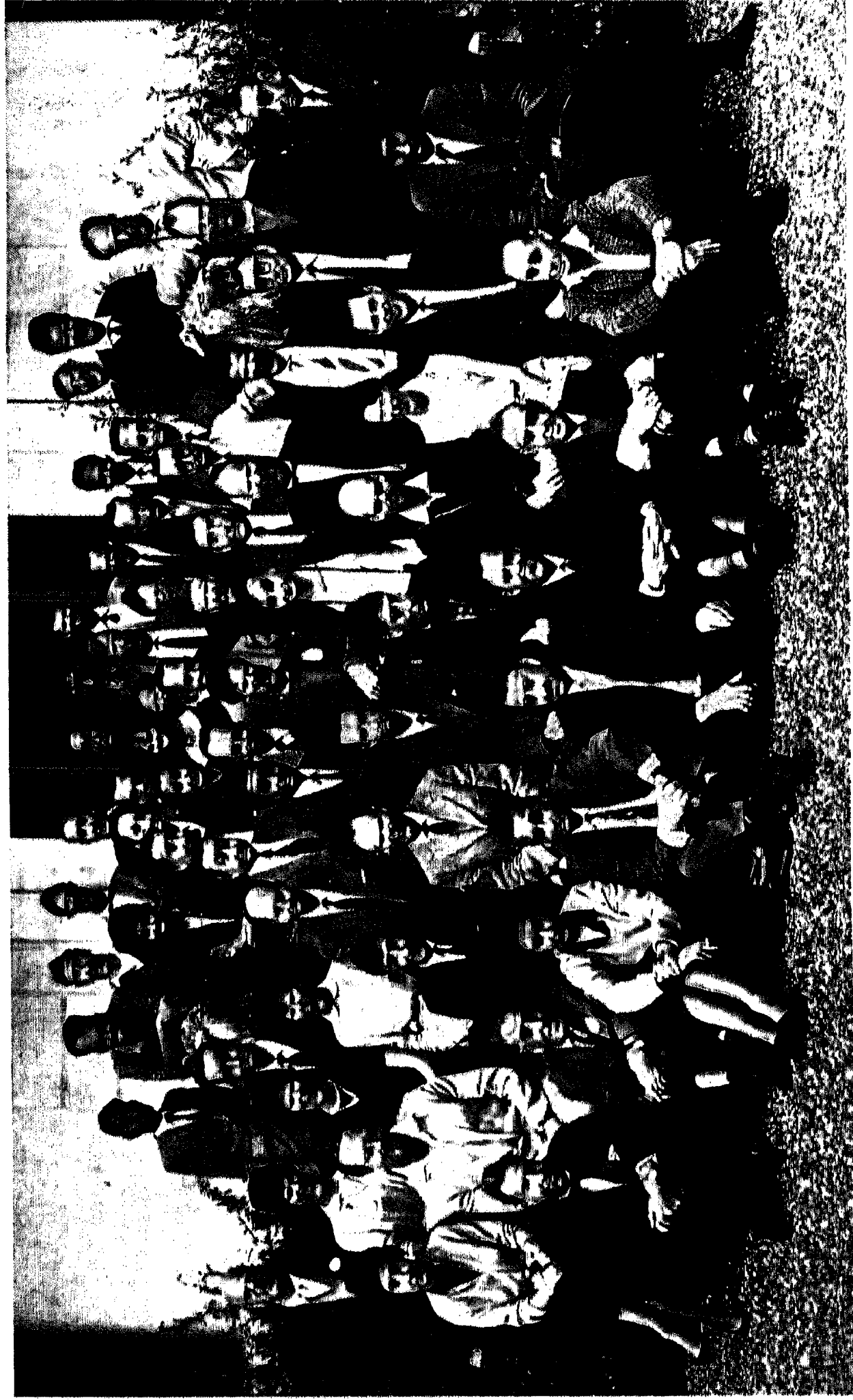
40. A. Bewick
41. H. Oeschger
42. H. Debrunner
43. J. Geiss
44. G. Germanà
45. L. Navone
46. J. H. Reid
47. L. Scatsi
48. M. Galli
49. D. E. Page
50. H. Winzler
51. C. J. Waddington
52. Sig.ra Waddington

53. G. H. Ludwig
54. N. C. Wichramasinghe
55. E. A. Brunberg
56. F. G. Houernians
57. C. Dilworth
58. B. Peters
59. J. Labeyrie
60. E. B. Dorling
61. J. M. Greenberg
62. B. Dayton

SOCIETÀ ITALIANA DI FISICA

SCUOLA INTERNAZIONALE DI FISICA «E. FERMI»

XIX CORSO - VARENNA SUL LAGO DI COMO - VILLA MONASTERO - 23 Maggio - 3 Giugno 1961



The Surface of the Sun.

M. HACK

Osservatorio di Merate

1. — Introduction.

Most of the disturbances occurring in the solar atmosphere, which are strictly connected with emission of X-rays and corpuscular radiation from the sun probably find their initial cause in the convective zone underlying the photosphere. For this reason, before starting to speak about the solar atmosphere and its disturbances we must give a brief account of the main properties of the solar surface, where evidence of the convective zone is apparent.

The sun is a gaseous sphere mainly composed of hydrogen (H is 85 % in number of atoms, He 15 %, for the other atoms we have $H/\text{metals} \simeq 10^4$); its radius is equal to 700 000 km and its mass is $2 \cdot 10^{33}$ g. Temperature, density and pressure steadily decrease from the centre to the surface. This sphere is in hydrostatic equilibrium under its own gravitation; this means that the pressure exerted by a column of gas lying below any layer is counter-balanced by the weight of the column of gas above the layer. The energy distribution of the continuous spectrum of the sun is very similar to that of a black body at the temperature of 5 700 °K. From the whole surface of the sun $4 \cdot 10^{33}$ erg s⁻¹ are emitted. The luminosity of the sun is maintained by production of energy by means of nuclear reactions occurring in the core. At the temperature and density of the solar core ($T_c = 15 \cdot 10^6$ °K, $\rho_c = 90$ g cm⁻³) the main nuclear reaction chain is the proton-proton cycle.

We must explain what we mean by solar surface, since this gaseous sphere does not have a well defined boundary.

We call the surface, or the photosphere, a layer dense enough to emit continuous radiation and to show a considerable brightness, but still sufficiently transparent to allow a direct investigation of the same layer. In other words the surface is the whole of layers from which we receive an appreciable percentage of radiation; its depth is about 300 km; radiation from layers deeper than 300 km is almost completely absorbed by the overlying layers. The opacity of the photosphere is mainly due to the absorption of negative

hydrogen H^- , and also to bound-free and free-free transitions of electrons moving in the fields of ions. The small depth of the photosphere, which is less than $1/2000$ of the solar radius, and which is seen from the earth under an angle of $0.5''$ explains why looking at the sun we have the impression of seeing a solid well bounded surface.

A direct proof of the semitransparency of the photosphere is given by the existence of limb darkening. The decrease of brightness from the centre to the limb of the solar disk is due to the inward temperature increase of the photospheric layers. When we look at the centre of the disk our line of sight enters normally to the surface of the sun: so we receive radiation from layers on the average deeper and hotter. When we observe the limb our line of sight enters tangentially to the surface and we receive radiation only from the cooler layers of the top. The limb darkening is a function of the wavelength, and is more apparent in the violet than in the red.

The photosphere is not homogeneous. Pictures of the sun in white light show a granular structure, which consists of bright blobs on a dark background. The sizes of the granules vary from $0.3''$ to $2.5''$ with a mean value of $1.0''$, corresponding to 700 km. The best photographs of granulation are those obtained by SCHWARZSCHILD, ROGERSON and EVANS from a balloon at an altitude of 25 km, to reduce the disturbing effect due to the turbulence of our atmosphere. The intensity difference between granules and dark background corresponds to an average difference in temperature of between 120° and 230° . These temperature differences refer to the mean level from which the continuous radiation emerges, i.e. from an optical depth of ~ 1 (which corresponds to a depth of ~ 280 km). The zig-zag structure of the absorption lines of the solar spectrum is also a proof of this lack of homogeneity of the photosphere in the higher layers where the absorption lines originate ($\tau \simeq 0.1$ corresponds to a depth of ~ 190 km). This zig-zag structure is an evidence that the spectra of the individual granules are slightly shifted with respect to each other because of Doppler effects (radial velocity difference ~ 0.4 to 1.5 km/sec). The granulation is the evidence of the existence in the upper parts of the sun of a zone where convective currents are present. Theory also requests that in those parts of a star where the gas is partially neutral and partially ionized convective currents must develop.

Observations of KIEPENHEUER [1] suggest that the granules have a magnetic field of the order of 400 G. This kind of observation is very difficult because of the small dimensions of the granuli. The existence of a magnetic field of this order of magnitude is supported by theoretical arguments: in a turbulent plasma, electric currents are generated between regions of different electron densities and temperatures; these produce weak magnetic fields. These fields can be increased in intensity by the so-called « spaghetti process ». Since the electrical conductivity of the ionized gas is very high the lines of force

are «frozen» in the gas. The turbulent motions tangle the magnetic lines of force. The continual twisting of the lines of forces serves to increase the number of lines through unit cross-sectional area, and hence the intensity of the magnetic field. This process will continue until equipartition between the turbulent and the magnetic energy density is reached: $H^2/8\pi = \frac{1}{2}\rho v^2$. From the solar models the density at the top of the granulation zone is about $3 \cdot 10^{-7} \text{ g cm}^{-3}$. The curve of growth gives for the turbulence $v \sim 1.5 \text{ km s}^{-1}$. It follows that $H = 290 \text{ G}$, a value in good agreement with the observation.

2. - Phenomena of solar activity on the solar surface.

The appearance of spots, changing in number and shape with time on the bright surface of the sun, is a well known phenomenon since the time of Galileo, who first observed it. Spots appear generally between $|5^\circ|$ and $|40^\circ|$ of latitude. None has been observed at latitudes higher than $|50^\circ|$. They appear either isolated or, more often, in groups. A typical spot consists of a dark umbra (temperature of 4000°K , to be compared with the 6000°K of the surrounding photosphere) surrounded by a penumbra. Zeeman splitting of the absorption lines is evident in the spectra of the spots. It is found that the magnetic field of the spots is mostly perpendicular to the surface, and the intensity is of the order of some hundred to some thousand gauss. There is a strict correlation between the spot and the intensity of the magnetic field, which can be represented by the formula [2]

$$H = 3700[A/(A + 66)] \text{ G},$$

where A is the area of the spot, in 10^{-6} of the visible hemisphere (so for A included between 10 to 400 we have H between 480 and 3200 G). Observation of Doppler shift of the lines in the spot spectra shows that gas of deeper layers of the photosphere is streaming out of the spot, while gas of high atmospheric layers is streaming into the spot (Evershed effect).

Spot groups or single spots always appear in a facula.

Faculae are bright structures, similar to clouds, slightly brighter than the surrounding photosphere. If there are no spots without faculae, the reverse is not true, and often small faculae appear without spots. Faculae have weak magnetic fields, of the order of a fraction of a gauss to several gauss.

Theory must try to explain the diverse features of the solar activity at the surface: appearance of spots and faculae, low temperature of the spots and high temperature of the faculae, magnetic fields which are so strong in the spots, the often observed bipolarity of the spot groups, the 11 year cycle and the inversion of polarity which happens at each cycle so that the full solar cycle is in fact 22 years.

No completely satisfactory theory has been proposed until now. On subject I will point out only this: until 20 years ago it was thought that spots were places of large convection, where the low temperature follows from the adiabatic expansion of the ascending gas. Now we know that the strong magnetic fields suppress convective motions and tend to bind together neighbouring masses of gas inhibiting flow across the magnetic lines of force. Rising gas cannot therefore flow horizontally and descend again to complete the convection cycle. So spots are regions of calm in a turbulent atmosphere.

Since convection is lacking, energy flow occurs only by radiation; hence it is suggested that spots are cooler because convective energy transport is inhibited in the presence of magnetism.

In the regions just outside the spots convection must be greater to make up the diminished heat flow over the spot areas. This could explain the brightness of the faculae which surround the spots.

We can compute that no convection will occur above the depth where the magnetic energy density is of the same order as the energy of turbulent motions. For the high values of the density which we find in the interior of the sun, the field H becomes less than $v\sqrt{4\pi\rho}$ so that in the deep layers convection will not be prevented. Apparently, by local violent convection, some of the field lines of the interior are brought to the surface, forming a spot or a group of spots.

3. — The solar atmosphere.

It is difficult to give a very short account of all the complex properties of the solar atmosphere. I will try to say briefly:

- a) What we mean by chromosphere and corona.
- b) Their main aspects derived by optical observations.
- c) A very short account of the theories explaining the support and heating of the solar atmosphere.
- d) What are the components of radio emission of the quiet sun.
- e) What are the components of radio emission of the active sun and their correlation with optical disturbances.

The layers overlying the photosphere are called the solar atmosphere. They include the chromosphere, or low-atmosphere, extending from the photosphere to 6 000 km of height, and the corona, a very rarefied envelope of ionized gas extending probably as far as the earth or further.

A model for the chromosphere and the corona has been computed

WOOLLEY and ALLEN; it explains fairly well the optical and radio spectrum of the chromosphere and corona, in absence of perturbation (see Table I).

TABLE I. - *Models of Woolley and Allen.*

h (10^3 km)	T ($^{\circ}\text{K}$)	$\log N_e$ (cm^{-3})
4	5 040	9,4
6	6 300	10,4
6,001	14 500	10,2
6,01	36 000	9,8
6,1	91 000	9,4
7	230 000	9,0
10	400 000	8,8
20	660 000	8,6
30	800 000	8,5
40	950 000	8,4
70	—	8,25
700	—	6,50

4. - The chromosphere.

The chromosphere has a spicular composition. Each spicule has a thickness of about 500 km; they are more or less perpendicular to the solar limb, and grow above the general level up 10 000 to 20 000 km and disappear in the course of some minutes.

The spectrum of these spicules show that they have temperatures of 10 000 to 20 000 $^{\circ}\text{K}$; they are hence cold elements in a hot corona, having a temperature of about 500 000 $^{\circ}\text{K}$.

The density in the spiculae is about 1000 times higher than in the surrounding corona.

5. - Prominences.

Prominences are visible as bright structure extending beyond the sun's limb, in the corona. When viewed on the disk they take the name of filaments. Typical dimensions are 200 000 km in length, 50 000 km in height, and 10 000 km in thickness. They are probably supported by magnetic fields; when an ionized cloud of gas falls in such a field the turbulent motions are stopped. The support is stable at these points where the magnetic field is horizontal. A quiescent prominence may become active, activation may be

due to flares. The equilibrium is perturbed, and the prominence moves along the magnetic lines of force and may disappear suddenly.

6. - Flares.

Flares are sudden brightenings of the faculae fields. They are mostly observed in monochromatic light (H_α); often they are observed several times in a day.

Only 17 cases are recorded to have been observed in white light. The areas are included between 100 and 1200 in units of 10^{-6} of the visible hemisphere.

They originate near quickly developing spots. Observations of limb flares show that they reach heights of 7000 to 16000 km on the surface.

Often the flare is surmounted by an ascending eruptive prominence (surge). Several concomitant phenomena occur together with a flare *i.e.* radiobursts and emissions of cosmic rays and X-radiation.

Flares are classified according to their area. They are classified in three classes of increasing importance: 1, 2 and 3. Subdivisions are 1^- and 3^+ (the first having area respectively less than 100 and larger than 1200. The area is expressed in 10^{-6} of the visible hemisphere). The light curve of the flare is steeper in the ascending part than in the descending part. Flares are probably regions where the temperature is the same as in the surrounding parts, but the electron density becomes 10^3 to 10^5 higher than in the surrounding part of the chromosphere and the corona. (N_e flare $\sim 10^{13} \text{ cm}^{-3}$, N_e chr. $\sim 10^{11}$, N_e cor $\sim 10^8$).

7. - The corona.

The corona is so faint compared with the brightness of the sky and with the solar light scattered by the optical parts of a telescope that it is impossible to observe it except during an eclipse. Only the coronagraphic equipment allows the study of the brightest parts of the corona outside an eclipse. The shape of the corona is a function of the 11 year cycle. Typical features are the streamers and the polar plumes.

The spectrum of the corona is composed: a) of a continuous spectrum due to solar light scattered by the coronal electrons (corona K); this fraction of coronal light is polarized; b) of a continuous spectrum due to solar light scattered by dust in interplanetary space (corona F). This light is probably unpolarized; c) monochromatic emission by Ca, Fe, Ni, Al ions, 9 to 14 times ionized. The ionization is due to collisions. The electron temperature of

corona in the regions where the yellow line of Ca XV at $5694 \text{ \AA}$ ($I \cdot P = 814 \text{ eV}$) appears may reach $5 \cdot 10^6 \text{ }^\circ\text{K}$.

8. – Support and heating of the chromosphere and corona.

How is the chromosphere supported? The great scale height is due to the intense turbulent motions having between 2000 and 4000 km an average value of 15 km/s.

Why has the chromosphere so high a temperature? The source of heating is probably the turbulence of the photosphere which in turn originates from the lower convective motions. The main part of the convective energy is dissipated in the convection zone but a small fraction of it moves upward in the form of travelling sound waves. The heating of the corona is explained in a similar way. The flow of mechanical energy coming from the photosphere, though dissipating part in the chromosphere, is still important when it arrives in the corona. We can compute that at a height of 6000 km there is equality between dissipation of mechanical energy and radiation. Higher up the temperature increases because the emission of radiation becomes relatively less important. The radiation of the corona is the order of $5 \cdot 10^3 \text{ erg cm}^{-2} \text{ s}^{-1}$, whereas the inflow of mechanical energy is about a factor of 10^2 greater.

This explains the sharp increase of temperature which is observed at 6000 km above the photosphere.

A part of the mechanical energy absorbed by the corona is lost by «evaporation». This evaporation must be important because the thermal velocity of the coronal gases is close to the velocity of escape. This proves that the hot coronal regions must necessarily be a source of particle emission.

Probably a mass of the same order as that evaporated is accreted from the interplanetary dust.

9. – Radio emissions from the sun.

Radio astronomical observations of the chromosphere and the corona give much information on the physical structure of these layers and their perturbations. In fact, as for all the ionized gas, they are optically thick for the radio waves, and hence good radio emitters, while in the optical dominion they are transparent and very poor emitters.

10. – Background radiation.

The solar radio emission is composed of a permanent part due to thermal radiation by the chromosphere and the corona; it is due to free-free transition of electrons in the field of ions, mainly protons. (The contribution of the

photosphere is completely negligible as a consequence of its low temperature.

It varies slowly in the course of the solar cycle.

This is called background radiation. Moreover we observe a transient component, presumably due to centres of activity, which consists of: slowly varying component, noise storms, and outbursts of several types.

The intensity of the quiet sun is obtained by eliminating the transient components. The computations of the emergent radiation for a chromospheric and coronal model offer no difficulties, and a comparison between observed and computed spectrum of the background radiation allows us to derive temperature and electron densities as a function of the height.

We remember that the refraction index μ of an ionized gas is equal $\sqrt{1 - (f_c/f)^2}$, where f_c in the mks-system is given by $f_c = 9\sqrt{N_e}$. For $f > f_c$ the radio waves propagation is normal. For $f < f_c$ μ is imaginary, the wave cannot propagate.

Going up in the corona, the electron density decreases, f_c decreases to zero and the propagation of lower and lower frequency waves becomes possible. For example wave-lengths shorter than 3 cm can propagate for the whole chromosphere and corona, while waves longer than 10 m can propagate only in the high-corona. Hence 3 cm radiation comes from the whole atmosphere but since the absorption coefficient of H II is given by $K \sim N_e/\mu f^2 T^{\frac{1}{2}}$ the chromosphere emission is more important than that of the corona. In fact observations show that the temperature of the disk at centimetric wave-length is that corresponding to the lower chromospheric layers, which are the main contributors of the emission; the temperature at metric wave-lengths is that corresponding to the coronal layers (for temperature we mean the equivalent temperature of the disk; i.e. the temperature of a black body of the same angular size of the sun which gives the same flux density which is observed for the sun). The angular size of the sun increases with the wave-length, and corresponds to the height of the different layers which are the main contributors to the emission.

11. — The slowly varying component.

The period of variation is about 27 days (which is the synodical rotation period). There is a strict correlation between sun spot area and flux density. The curve extrapolated to zero sunspot area gives the background radiation level. The origin of the slowly varying component is probably thermal. Since its spectrum extends only from 2 to 60 cm it is thought that the layers where this radiation arises are in the chromosphere and in the low corona.

It probably originates in the coronal condensations. These condensations are found in the region over a spot group; the electron density in the condensations is higher than in the near regions.

The temperatures of the coronal condensations were believed to be higher than in the nearby regions. Recent observations by CHRISTIANSEN and MATHEWSON with an interferometer show that this is not case and the temperatures are equal to the surrounding parts.

The temperature and density of the condensations permit us to compute the expected radiation by thermal mechanism; this is in agreement with the observed radiation; so one is justified in thinking that this component is thermal.

12. - Noise storms and type I bursts.

They appear on areas overlying a big group of spots.

Their spectrum is extended between 1 m and 15 m. The equivalent temperature attains 10^9 °K and the emission mechanism must be nonthermal. Radio storms are in correlation only with a spot group the area of which is larger than $400 \cdot 10^{-6}$ of the visible hemisphere; this has been interpreted as due to the fact that radio-storms can arise only when the magnetic field of the spot is

TABLE II. - *Non-thermal radio phenomena.*

Class	Duration	Polarization	T_A	Associated optical features	Spectrum	Dynamic spectrum	Corresponding speed km/s
Type I or storm bursts	Individual lifetime ≤ 1 s; the noise storm: hours or days	Mainly polarized	10^9	Big spot groups ($A > 400 \cdot 10^{-6}$)	1 to 15 m		
Type II	(1 ÷ 10) min	Variable	$> 10^{10}$	Flare	cm, dm, m	$df/dT \sim \sim -0,1$ MHz	10^3
Type III	$\leq$ some seconds	Variable	$> 10^{11}$	Small flare	cm, dm, m	$df/dT \sim -100$ MHz	10^5
Type U	$\leq$ some seconds	Variable	$> 10^{11}$	Small flare	cm, dm, m	df/dT from + 100 to - 100 MHz	10^6
Type IV	Hour	Circular	10^{11}	Flare	cm, dm, m		
Type V	Minutes	Circular	$> 10^{11}$	Flare	cm, dm, m		

sufficiently strong to reach a height of 60 000 km (corresponding to the metric waves, emitted by the storms). The spectra of the bursts correlated with the storms are characterised by having a very short wave-length band (few MHz).

So the storms are observed at the same time at metric and decametric wave-length, but there is a big difference in the details since there is no correlation at all between bursts even at closely similar wave-lengths.

13. – Type II and III, IV and V bursts.

Type II and III bursts consist in sudden increases of solar radiation corresponding to apparent temperatures of the order of 10^{10} or 10^{11} °K; since they cover only small areas of the solar disk, the brightness temperature is much higher. But the most interesting characteristic of these bursts is that they do not occur at the same time at the different frequencies, but first at shorter wave-lengths and later at longer wave-lengths. The « dynamic spectra » of type II bursts show a frequency drift $df/dt \simeq -0.1 \text{ MHz s}^{-1}$ to -1 MHz s^{-1} .

Type III burst show a frequency drift $df/dt \simeq -100 \text{ MHz s}^{-1}$.

For this reason they are also called respectively slow drift and fast drift burst. Remembering that an increasing wave-length of the radiation emitted corresponds to an increasing height in the solar atmosphere this frequency drift corresponds to a velocity of propagation of the disturbance of some 100 km/s to 1000 km/s for the slow drift burst and of about 100 000 km/s for the fast drift bursts.

Type II bursts last about 10 minutes and occur in the postmaximum phase of certain large flares. Polarization is usually random.

Type III bursts are common events lasting about 10 seconds. They occur in groups near the start of flares. Polarization is usually weak or random. Type IV bursts are associated with type II. They are rare events lasting 10 minutes to hours and start late in the life of certain flares, following a type II burst. Their spectrum is a featureless continuum. Polarization is circular.

Type V bursts are associated with type III. They last about 1 minute and immediately follow certain type III bursts. Their spectrum is a featureless continuum. Polarization is circular.

U bursts: are a variety of type III bursts; the frequency drift starts normally, becomes retarded and then completely reverse, so the dynamic spectrum resembles an inverted U.

More than 50 % of the small flares are followed by type III bursts, but very few or none of the larger flares. The contrary is true for type II: none of the small flares but almost 30 % or more of the larger flares are followed by type II bursts.

The probability for a type II burst to be followed by a type IV is $10 \div 20\%$. If the flare is of importance 3 the probability becomes 30% . 50% of type II

TABLE III. - (From J. P. WILD: *Rendiconti S.I.F. Corso XII* (1959)).

Flare importance		1	2	3
Type III	20 .25	50	?	?
Type II	0	2	7	31

bursts and a few % of type III bursts show duplication of their spectral features at twice the frequency; *i.e.* the same bursts radiates at a fundamental frequency and at its second harmonic.

14. - Interpretation of the phenomena correlated with bursts of type II, III IV and V.

The existence of harmonic in type III and II bursts indicates that emission is generated by electrons oscillating at a proper frequency. Since the critical frequency (or plasma frequency $f_c = 9 \sqrt{N_e}$ MHz, with N_e in m^{-3}) is continuously varying in the solar atmosphere with increasing height, the view that type II and type III bursts originate in plasma oscillation has been accepted. According to the plasma hypothesis the systematic frequency drift is interpreted as an outward motion of a disturbance which excites plasma oscillation in the surrounding coronal gas; the greater the height of the disturbance, the lower the electron density and hence the lower the frequency emitted. Knowledge of the electron density as a function of the height allows us to derive the radial component of the velocity of the source. A direct verification of this hypothesis has been given by WILD, SHERIDAN and NEYLAN [3], observing interferometrically about 100 type III bursts at the limb of the sun. From the direct observations of the tangential motions of the disturbance it was found that the regions of origin are at much greater height than expected from the application of coronal models of electron density. The model required to fit the observation needed electron densities to be increased by a factor of about 10. However the required model agrees closely with that deduced optically for typical coronal streamers.

The wider separations of plasma levels indicated by the interferometer observations at the limb, lead to a revision of the velocity of type III disturbance. The mean velocity is $\sim 0.5 c$ and there is a sharp low-velocity cut-off at $0.2 c$.

Corrections must be applied also to the velocity of type II bursts which are found to be of 1000 to 1500 km/s.

The type V bursts lack the features such as discrete bands, frequency drift

and harmonics indicative of plasma emissions. It has been suggested that synchrotron radiation from electrons spiralling in a magnetic field can explain type V and type IV bursts-radiation. WILD [4] computes that at least 10^{33} electrons with energies of 2 MeV or greater are required to explain the observed emission. A magnetic field ≤ 1.5 G is admitted for the upper parts of the corona. Since observations show that the emitting region attains large sizes ($5 \cdot 10^5$ km) this total number of electrons corresponds to an electron density less than 30 cm^{-3} , which is 10^{-5} the electron density of the corona at 700 000 km from the limb. Hence this density of relativistic electrons is acceptable. The observed polarization of radiation from type V and IV bursts is another point in favour of the hypothesis of synchrotron radiation. WILD [4] suggests the following sequence of the events: a flare is originated by an explosion near the photosphere; this explosion causes a sudden increase in the emission of H_{α} light (the flare); ejects a stream of relativistic electrons which escapes immediately into the outer layers of the corona and beyond (III and V type bursts) and sets up a shock-wave which carries with it further relativistic electrons (II and IV types bursts). The physical nature of this postulated explosion is an unsolved problem. Almost 10^{33} relativistic electrons are produced during a flare; the particles follow a spiral path along the lines of force, which are materialized by the coronal streamers and excite the coronal plasma, causing type III bursts. The same relativistic electrons radiate synchrotron radiation causing type V bursts. In several cases the particles escape in the interplanetary space, but in other cases they come back following a line of force; so we have the U bursts. The shock wave originated by the same explosion moving in the coronal plasma at a velocity of about 1000 km/s (about 10 times the sound velocity in the corona) excited the plasma and causes type II bursts. The magnetic field and relativistic electrons carried by the shock-wave originate type IV bursts.

Since the radio emission from flares is correlated to the ejection of solar particles, we must expect that cosmic ray increases and geomagnetic storms are correlated with the appearance of flares.

The occasional large increases in the flux of cosmic radiation observed on the earth following flares start 10 to 30 minutes after the start of the flare, despite the relativistic velocities of the particles. It has been suggested that their emission from the sun must be delayed. DENISSE, BOISCHOT and PICK-GUTMANN [5] find that with a few exceptions, these flares are of importance 3 and 3^+ (the most important observed flares, according to the symbols used by the astrophysicists) and are followed by the most intense type IV radio bursts. Since all the increases in cosmic radiation correspond to type IV bursts, they suggest that type IV sources are responsible for ejecting cosmic rays. The delay between flare and cosmic rays increase is explained by OBAYASHI and HAKURA [6] assuming that the ejected clouds would trap the magnetic

fields of sunspots or the corona and travel outward through the solar atmosphere into interplanetary space. High-energy particles may be trapped partly in such magnetic clouds during the earlier stage of their passage through interplanetary space. However, as the cloud undergoes an expansion and its magnetic field is diluted, then high energy particles may gradually escape from the cloud and reach the earth.

The recognition of sudden commencements of terrestrial magnetic storms 1÷3 days after large flares suggests a relationship with type II outbursts. No significant correlation with type III bursts was found. The correlation between geomagnetic storms and IV type bursts is close. These results suggest that the presence of a type IV burst is an indication of the amount of matter transported from the flare by the second-phase disturbance.

DE JAGER suggests [7] that solar X-rays at short waves lengths (~ 1 Å) and type III radio bursts are both caused by similar streams of fast electrons. The X-radiation would be synchrotron radiation.

This hypothesis has been confirmed by observation of WINCKLER, MAY and MASLEY of a strong X-ray burst (11 August 1960) which coincides exactly in time with a burst of type III observed by MAXWELL *et al.*

REFERENCES

- [1] K. O. KIEPENHEUER: *Astraph. Journ.*, **117**, 447 (1953).
- [2] J. HOUTGAST and A. VAN SLUITERS: *BAN*, **10**, 325 (1948).
- [3] J. P. WILD, K. V. SHERIDAN and A. A. NEYLAN: Paris Symposium on Radio Astronomy, p. 176.
- [4] J. P. WILD: *Rendiconti S.I.F., Corso XII* (1959), p. 296.
- [5] J. F. DENISSE, A. BOISCHOT and M. PICK-GUTMANN: *Space Research. Proc. of the First Int. Symposium* (Amsterdam, 1960), p. 637.
- [6] T. ODAYASHI and Y. HAKURA: *Space Research*, p. 665.
- [7] C. DE JAGER: *Space Research* (1960), p. 628 and *X Colloque de Liège* (1960), p. 653.

An Optical Phenomenon Associated with Type IV Solar Radio Emission.

J. H. REID

Dunsink Observatory - Dublin

This new phenomenon, observed by hydrogen H_α -light, is associated with some Class 3 and 3+ solar flares and appears as a dark halo surrounding the bright emission flare filaments. We propose to call this effect the «flare nimbus» [1].

So far we have observations of five cases:

1959 July 16

1960 April 1

June 1

November 10

November 12

The first and last of these flares generated high energy protons which resulted in secondary cosmic ray effects at ground level, the other three were followed by polar cap absorption events and all five flares had emission of radio noise of Type IV.

The flare nimbus first appears near the time of the maximum intensity of the flare filaments as seen by hydrogen light, and in only one case (June 1) was it seen before the maximum. The most conspicuous phase of the nimbus appears about 30 minutes *after* the flare maximum and its dimensions lie in the range $(2 \div 4) \cdot 10^5$ km or $4.5' \div 9'$. The duration of the nimbus is $> (1 \div 2)$ hours.

The continuum radiation of Type IV usually begins near the time of the flare maximum (observed optically) and has a duration of the order of several hours. It reaches maximum intensity some $(20 \div 30)$ minutes after the start and the size of the source on the sun, as measured at 169 MHz with the Nangay interferometer, is of the order of $7' \div 12'$.

Thus, the flare nimbus and the Type IV radiation source have comparable characteristics and we believe that the nimbus may be regarded as the optical counterpart of the cloud of relativistic electrons postulated by BOISCHOT and DENISSE [2] as the cause of the continuum radio emission.

It is not very obvious why these high-energy electrons should have any visible effects at the chromospheric level. The electrons, however, are not stationary and if we assume that each is moving in a helical trajectory along lines of force they will perform oscillatory movements between the reflection points in the magnetic field.

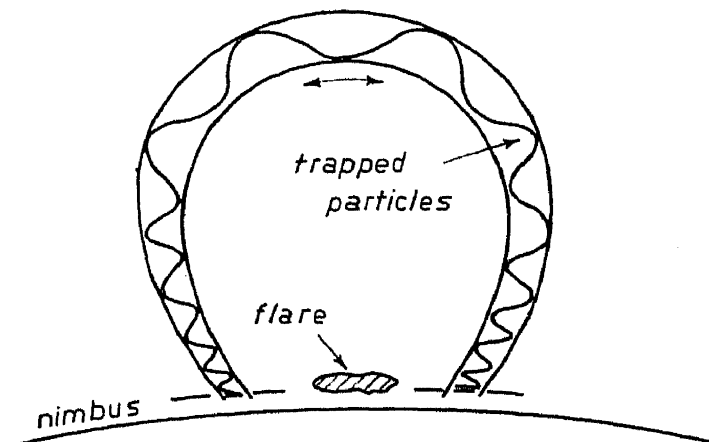


Fig. 1.

A typical magnetic loop in the coronal region is shown in Fig. 1 and such a region may act as a trap for these high-energy particles during their lifetime, which is of the order of a few hours.

Energy decay through collisions will be most rapid at the reflection points in the chromosphere and the flare will be surrounded by an «auroral zone» in which the collisions will be most numerous. Thus the absorption of light may be caused by the excitation of neutral hydrogen atoms in the second quantum level.

Note added in proof.

The results of a recent photometric investigation [3] of the nimbus phenomenon suggest that it is not an absorption effect as suggested above, but arises from the brighter parts of the striation pattern surrounding the flare returning to the intensity level of the darker background, due perhaps to the extraction of magnetic energy from the region. It is believed that flares producing Type IV radio emission may eject the active region magnetic field into space [4] and the relation of the nimbus to this type of radio emission can thus be accounted for.

REFERENCES

- [1] M. A. ELLISON, S. M. P. MCKENNA and J. H. REID: *Month. Not. Roy. Astron. Soc.* (in press).
- [2] A. BOISCHOT and J. F. DENISSE: *Compt. Rend.*, **245**, 2194 (1957).
- [3] J. H. REID: *The Observatory* (in press), and a more detailed account is being prepared for publication later
- [4] J. P. WILD: *Journ. Phys. Soc. Japan*, **17**, Suppl. A-II, 249 (1962).

Motion of Cosmic Ray Particles in the Geomagnetic Field.

H. ELLIOT

Physics Department, Imperial College - London

A proper understanding of the motion of cosmic ray particles in the earth's magnetic field is important for a number of reasons. In the first place, we can use the earth as a magnetic spectrometer enabling us to measure the magnetic rigidity spectrum of the primary particles and to determine how this varies in time as a result of the various solar modulation processes. It also enables us to measure the spectrum of solar particles accelerated in flares and having momenta in the region 2 to 15 GeV/c. Again, in interpreting anisotropies in the radiation incident at the earth's surface, we need to know the relationship between the incident direction at the earth and the asymptotic direction outside the field as a function of particle momentum. Finally, the cosmic ray intensity changes which arise from distortion of the outer regions of the geomagnetic field provide us with indirect information about the nature of these field distortions. We can only expect to derive useful information in the latter case if we have a complete understanding of the effect of the internal parts of the earth's field, and I shall begin by considering this part of the field.

The magnetic potential due to sources inside the earth may be represented to any required degree of accuracy by means of spherical harmonic functions. Thus we may write the potential, V , in the form

$$(1) \quad V = \sum_{n=1}^n V_n \quad \text{where} \quad V_n = r_e \left(\frac{r_e}{r} \right)^{n+1} T_n$$

and r_e is the radius of the earth, r the distance from its centre, and T_n is a complicated function of latitude and longitude. $n=1$ corresponds to the axial dipole term, $n=2$ to the quadruple and so on. The greater part of the field arises from the dipole term which, according to the 1955 magnetic data, corresponds to a magnetic moment of $8.1 \cdot 10^{25} \text{ G m}^3$. The axis of this dipole cuts the earth's surface in the northern hemisphere at 78.2° N and 69° W .

A better representation of the field as a whole is obtained by displacing

the dipole to a position 433 km from the geocentre and at 15.3° N and 151.1° E. This representation takes into account part of the $n=2$ term neglected in the centred dipole approximation. All the early work on the motion of cosmic ray particles in the geomagnetic field was based on the dipole representation although, as we shall see, this is not in fact a particularly good approximation.

For a dipole field there are, in general, certain regions of space which are accessible to particles coming from infinity whereas other regions are inaccessible. For an axially symmetric system, that is to say a dipole field embedded in an isotropic flux of incident particles, the forbidden regions must also be axially symmetric, and it is therefore sufficient to consider the intersection of these regions with a meridian plane. In such a plane, the forbidden and allowed regions for particles approaching from infinity are specified by the « Stormer integral » of the equations of motion,

$$(2) \quad 2\gamma = R \cos \lambda \sin \theta + \frac{\cos^2 \lambda}{R}$$

where γ is a constant which is proportional to the impact parameter relative to the dipole axis of the incident particle at infinity, λ is the geomagnetic latitude, θ is the angle between the velocity vector of the particle and the meridian plane, and R is the radial distance from the dipole measured in Stormer units. (One Stormer unit is equal to $\sqrt{M/P}$, M being the dipole moment and P the particle rigidity given by $P = pc/Ze$, where p is the momentum, c the velocity of light, and Ze the particle charge.)

Since $\sin \theta$ must lie between ± 1 for a real particle, the boundaries between the allowed and forbidden regions are obtained by putting $|\sin \theta| = 1$ in eq. (2).

In general there are two allowed regions—the outer one extending to infinity and the inner one which contains the dipole. These two regions may or may not be connected, depending on the values of γ . For $\gamma < 1$ the regions connect, whereas for $\gamma > 1$ they do not. Figure 1 shows the appearance of the allowed and forbidden regions at their intersection with a meridian plane.

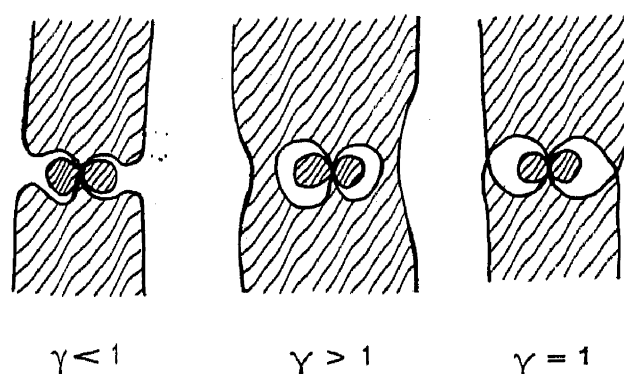


Fig. 1.— Allowed and forbidden regions in a meridian plane

It can be shown that the minimum rigidity which a particle must have to reach the earth's surface is given by putting $\gamma = 1$. At this value of γ the outer and inner regions are just beginning to connect. We are most frequently interested, in cosmic ray problems, in the minimum rigidity a particle needs to arrive at the earth's surface at vertical incidence, which corresponds to

$\sin \theta = 0$. Thus for this state of affairs eq. (2) becomes

$$2 = \frac{\cos^2 \lambda}{R}$$

where R is measured in Stormer units, *i.e.*,

$$R = \frac{r_e}{\sqrt{M/P_{\min}}},$$

r_e being the radius of the earth in cm; then

$$(3) \quad \frac{r_e}{\sqrt{M/P_{\min}}} = \frac{\cos^2 \lambda}{2} \quad \text{or} \quad P_{\min} = \frac{M}{4r_e^2} \cos^4 \lambda = 14.9 \cos^4 \lambda \text{ in GeV.}$$

LEMAITRE and VALLARTA have shown by the integration of individual trajectories that whilst particles with rigidity less than $P_{\min}$ cannot reach the earth at latitude λ it does not always follow that particles with rigidity $P_{\min}$ or greater can do so. There are in fact forbidden regions additional to those given by the Stormer relation and those produce an increase in the value of $P_{\min}$ over that expected from Stormer theory. Corrections for this have been calculated by KASPER so that the accurate values of $P_{\min}$ for a dipoles field are as follows.

	70	60	50	40	30	20	10	0
$P_{\min}(S+K)$	0.20	0.92	2.52+0.15	5.2+0.4	8.5+1.3	11.7+1.5	14.1+0.4	14.9

Whilst these values of $P_{\min}$ are accurate for a dipole field they are true for the earth only in so far as the geomagnetic field can be represented by a dipole to the exclusion of all the higher order terms in eq. (1), Nevertheless for many years they sufficed.

Discrepancies between the measured cosmic ray distribution over the surface of the earth and that to be expected for a dipole field appear to have been first noticed by JOHNSON in 1938, who pointed out that there appeared to be a correlation with the local field intensity. The introduction of nucleonic detectors about ten years ago led to much greater sensitivity in determining the dependence of the cosmic ray intensity on latitude and longitude, and the work of ROSE, FENTON, KATZMAN and SIMPSON revealed the existence of serious discrepancies. In particular, at 30° W longitude the minimum cosmic ray intensity occurred 9° N of the geomagnetic equator. At about the same

time WADDINGTON reported significant discrepancies between the α -particle intensities observed at balloon altitude at similar geomagnetic latitude over Europe and North America.

ROTHWELL and QUENBY, in 1957, showed that the intensity of the nucleonic component was closely related to the local values of the field parameters, indicating that the higher order terms in eq. (1) are of importance in determining the threshold rigidities for primary particles reaching the earth.

There is no simple analytical integral of the equations of motion when the non-dipole terms are taken into account, and the only rigorous solution of the problem involves the integration of a large number of trajectories in order to identify the asymptotic one corresponding to P_{min} . The effect of the high order terms is very great in some parts of the world. In the South African regional anomaly, for instance, the correction to the dipole threshold rigidities may amount to as much as 30 or 40%. Consequently a comparatively crude correction is much better than none at all. QUENBY and WEBBER have made an approximate calculation of the corrections involved, taking into account the spherical harmonic terms from $n=1$ to $n=6$.

Terms higher in the series than this have a negligible effect. Tables setting out these corrections were published by QUENBY and WEBBER in 1959 and some refinements of their calculations have since been carried out by WENK.

Figure 2 shows the lines of equal threshold rigidity calculated by WENK. These calculations, which are only approximate, are based on the assumption that the motion of the cosmic ray particles in the earth's field is, to a first approximation, the same as for a dipole field. In the equatorial region, the

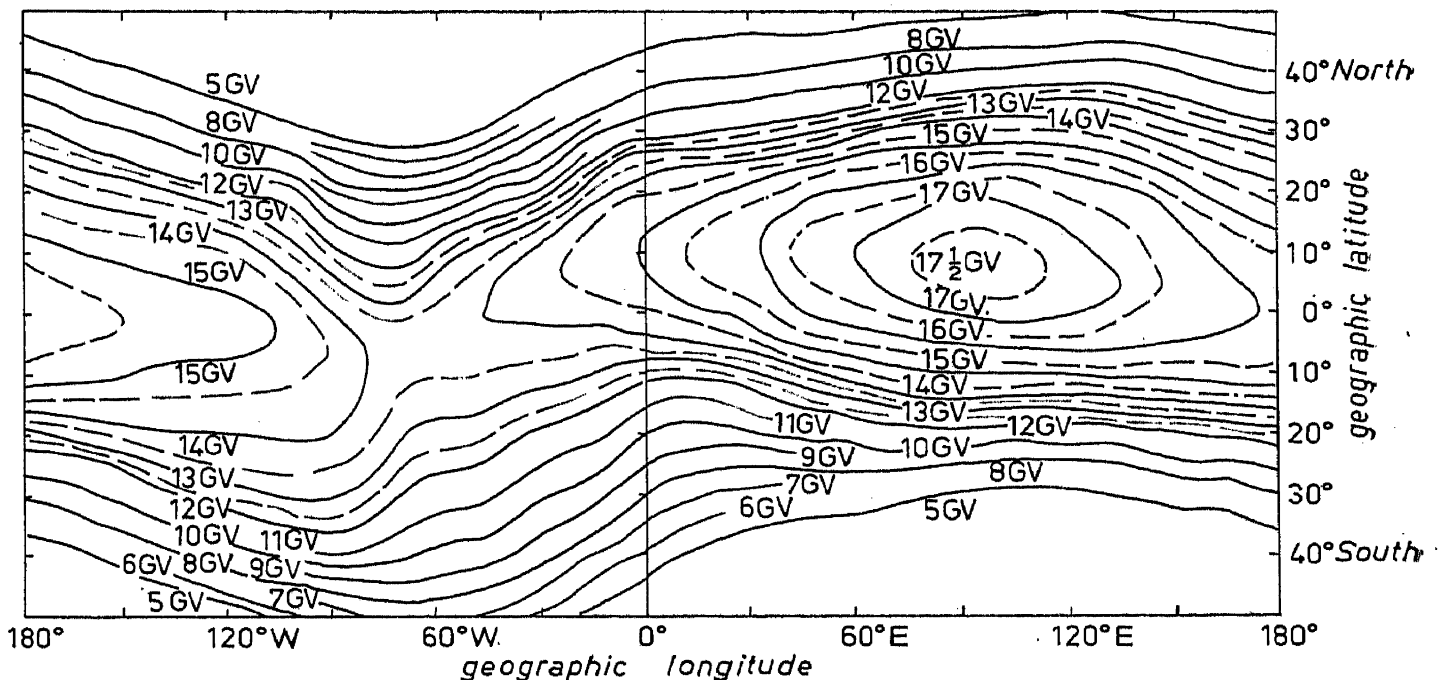


Fig. 2. — Lines of equal threshold rigidity (QUENBY and WENK).

higher order terms are taken into account by changing the dipole moment, whereas at high latitudes they result in a change in the geomagnetic latitude, the changes being deduced from Hultquist's computation of the end points of lines of force in the real field of the earth. The values of threshold rigidity obtained in this way are probably never wrong by more than 10% or so.

More recently KELLOGG has determined the threshold rigidities in the equatorial region by integrating trajectories in a field represented by the sum

of the first six terms. He has used the values of the threshold rigidity deduced in this way to compute the position of the cosmic ray equator.

Figure 3 shows a comparison between the cosmic ray equators deduced from the work of QUENBY and WEBBER and of KELLOGG and that determined experimentally by SIMPSON.

The calculations of QUENBY and WEBBER and WENK refer only to the Stormer thresholds and take no account of the penumbra. An empirical correction for the penumbra has been derived by QUENBY and WENK in collaboration with the Chicago group using the data from Simpson's airborne survey of the neutron intensity in the equatorial region. This has been done in the

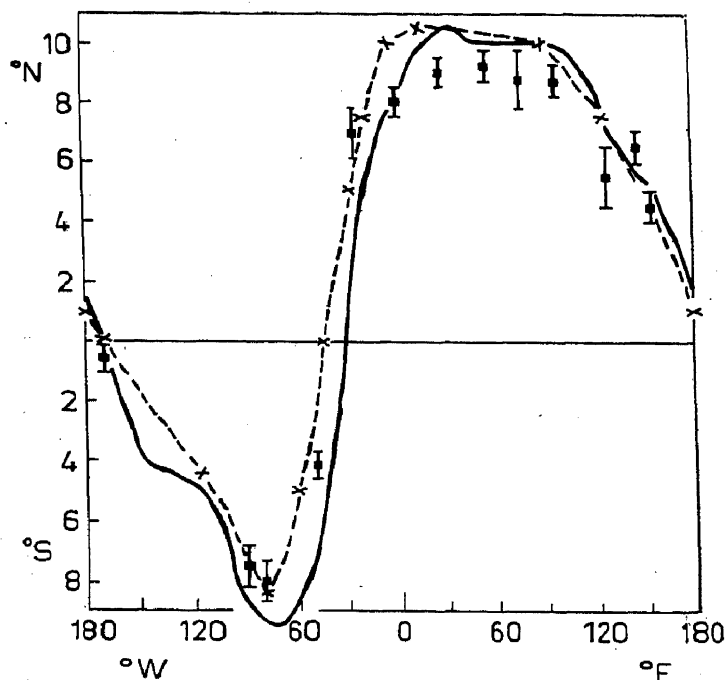


Fig. 3. — Comparison between observed and predicted positions of the cosmic ray « equator »: $\blacksquare$ SIMPSON *et al.*; — QUENBY and WEBBER; — x — KELLOGG (2° integration).

following way. Along the cosmic ray equator the effect of the penumbra is negligible but the threshold rigidity, and therefore the neutron intensity, depends upon longitude. Consequently, by comparing the latitude variation in the neutron intensity with the variation in intensity with longitude the change in threshold rigidity due to the penumbra can be evaluated. This correction has been included in the map shown in Fig. 2. Figure 4 shows a comparison between the empirical correction for the penumbra deduced in this way and the correction calculated by SCHWARTZ for a pure dipole field.

The discrepancies between positions of the geomagnetic and cosmic ray equators were at first incorrectly attributed to the distortion of the outer reaches of the dipole field by external current systems. There now exists, however, a considerable body of evidence which can only be interpreted in terms of such current systems.

The first definite evidence for an increase in the cosmic ray intensity at

sea level, due to the external deformation of the field, was obtained by Japanese workers. Using data for the magnetic storms of September 13, 1957 and February 11, 1958, they have shown that the anomalous increase in intensity observed during the main phase of these storms could be explained as a lowering of the geomagnetic threshold rigidities by the external current system. Changes in threshold rigidity of up to 1 GV are required to account for the observations.

Direct evidence for the arrival of solar flare particles at the times of magnetic storms with rigidities below those appropriate to the undisturbed field has been obtained by a number of workers, and the changes in cosmic noise absorption at high latitudes during these events also indicate severe distortion of the field.

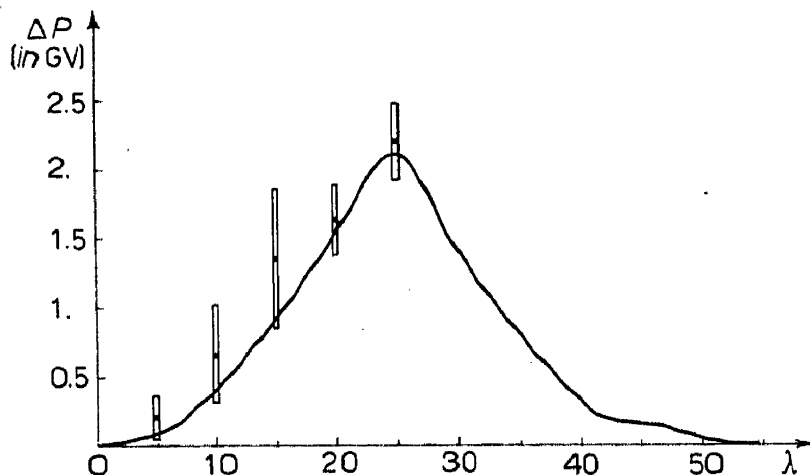


Fig. 4. — Penumbral correction to geomagnetic threshold rigidities.

The most important form of field distortion we have to consider—although perhaps not the only one—is that produced by a ring current generated by plasma trapped in the field. The existence of a ring current was postulated long ago to explain the motion of auroral displays towards the equator at times of magnetic disturbance and to account for the equatorial field changes during the main phase of a magnetic storm. Results strongly supporting the ring current hypothesis have recently been obtained by means of earth satellite and space probe magnetometer measurements. Data from Explorer VI and Pioneer V obtained by the group at Space Technology Laboratories are consistent with a current of the order $5 \cdot 10^6$ A flowing in a westerly direction at a distance about 10 earth radii out from the centre of the earth. The current system is supposed circular in cross-section. In order to produce the best fit with the measured field the radius of this circular cross-section has to be about three earth radii. This current system appears to be a permanent feature of the outer geomagnetic field.

The effect of an axially symmetric system on the geomagnetic threshold rigidities has been discussed by TREIMAN and by RAY. They have shown that the effect of such a current system is to lower the threshold rigidities below those which prevail for a purely dipole field.

Figure 5 is a diagram prepared by WEBBER showing the changes in threshold rigidity produced by three different ring current configurations according to Ray's calculations, which are based on a filamentary ring current. Also

shown on the same diagram are the changes to be expected if a uniform large scale field, oppositely directed to the earth's field at the equator, is superimposed externally. This situation would correspond physically to that where the earth is situated in a large scale interplanetary field with a direction perpendicular to the plane of

the ecliptic. The existence of a field of this kind has been inferred by SONETT from magnetic measurements in the space probe Pioneer V.

On the diagram, WEBBER has plotted a few points indicating the changes in geomagnetic threshold which he has deduced using data obtained during the solar proton/magnetic storm events of May 11 and July 14-17, 1959.

Figure 6 shows a comparison between calculated and observed threshold rigidities over North America. There appears to be a real indication that the observed thresholds are in general lower than those calculated on the basis

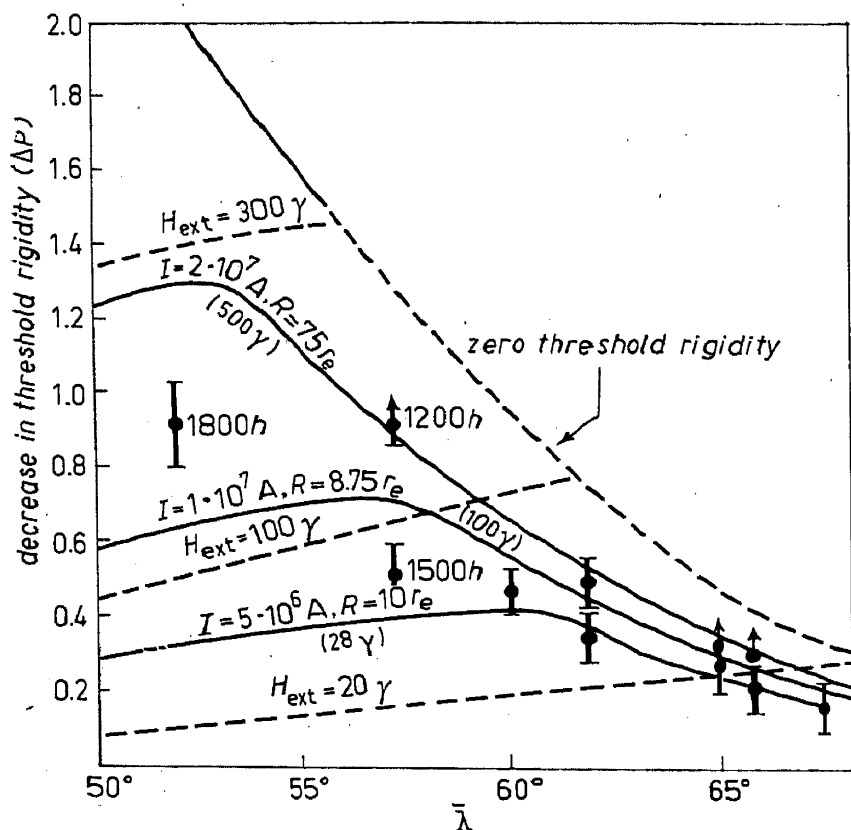


Fig. 5. - Changes in threshold rigidity produced by external current systems (after WEBBER).

of the internal field only. This situation would be compatible with the existence, as a normal state of affairs, of a ring current of the magnitude suggested by the Explorer VI and Pioneer V measurements.

Up to the present time the only calculations of the effects of external current systems have been carried out for axially symmetric fields for which a Stormer integral can be found, that is to say for an internal dipole field plus either a uniform or symmetrical external field. The computation of geomagnetic thresholds for complicated non-symmetrical fields is likely to be expensive in terms of machine time.

At Imperial College a programme of threshold determination has been started by C. J. BLAND using a magnetic scale model technique basically similar to that employed by MALMFORS and by BRUNBERG and DATNER for determining the asymptotic direction of cosmic ray primaries incident on the earth. The magnetic field of the terrella is generated primarily by means of coils, the regional anomaly field being produced by permanent magnets situated inside the terrella at suitable points about half-way between the centre

and the outside surface. In the case of the earth it is known that the anomaly fields can be represented by radial dipoles situated near the surface of the core. The cosmic rays are simulated by electrons fired outwards from the

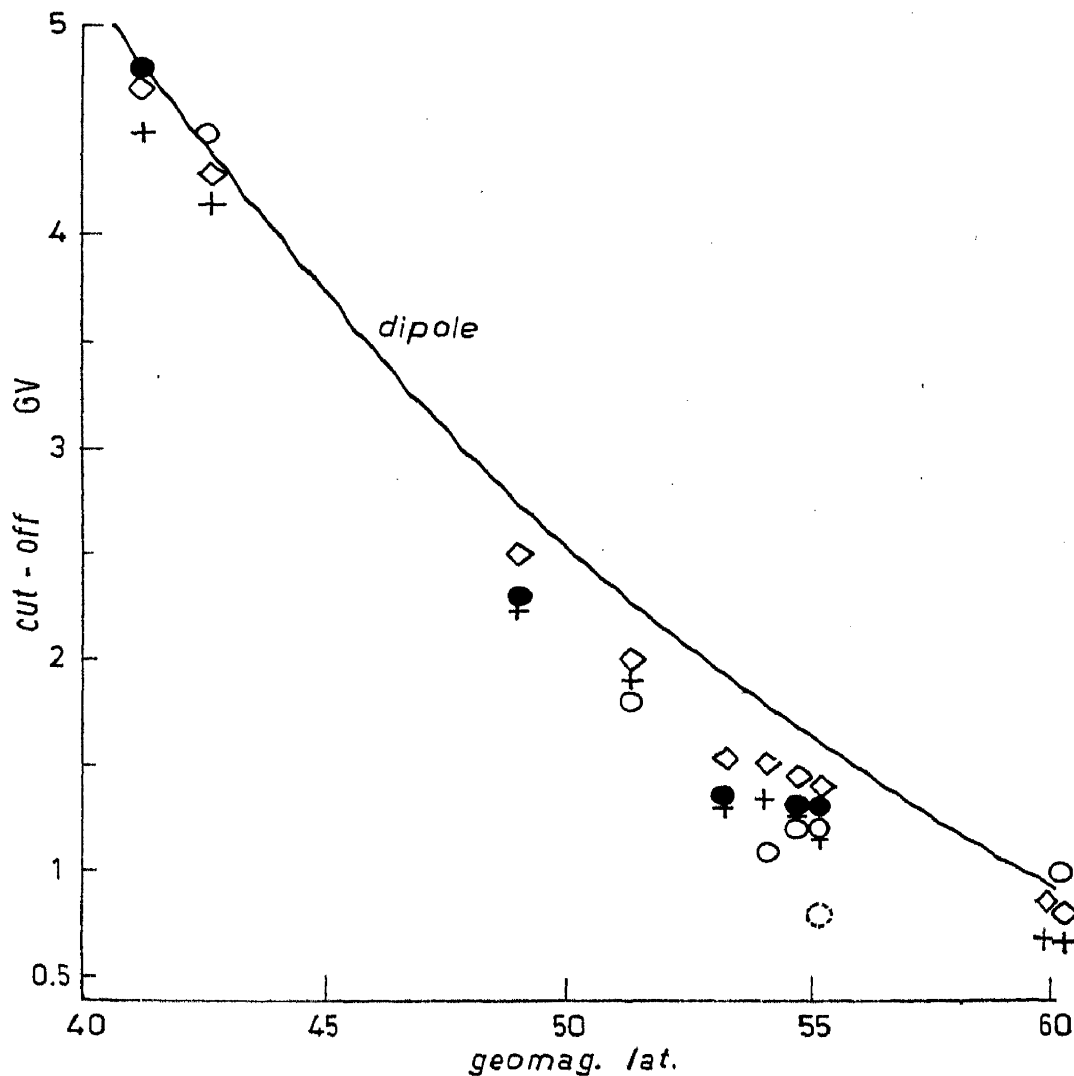


Fig. 6. — Comparison between observed and calculated threshold rigidities (after QUENBY): ● C-S (observed); ○ emulsion (observed); ◇ HULTQUIST; + QUENBY and WEBBER.

surface of the terrella with an appropriately scaled energy. The radius of the terrella is 10 cm, the surface field ~ 80 G, and the energy of the electrons in the neighbourhood of 200 eV.

Figure 7 shows the terrella positioned inside a vacuum tank which is cylindrical in shape having a diameter of 30 in. and length 3 ft. 3 in. This tank can be pumped down to a pressure of 10^{-6} mm Hg and the whole apparatus is situated in a field free region produced by the large coils shown in Figure 8.

In determining the geomagnetic threshold rigidities the electron gun is positioned at the appropriate latitude and longitude and the ratio of the gun current flowing to the terrella, i_{terrella} , to that flowing to the tank, i_{tank} , is

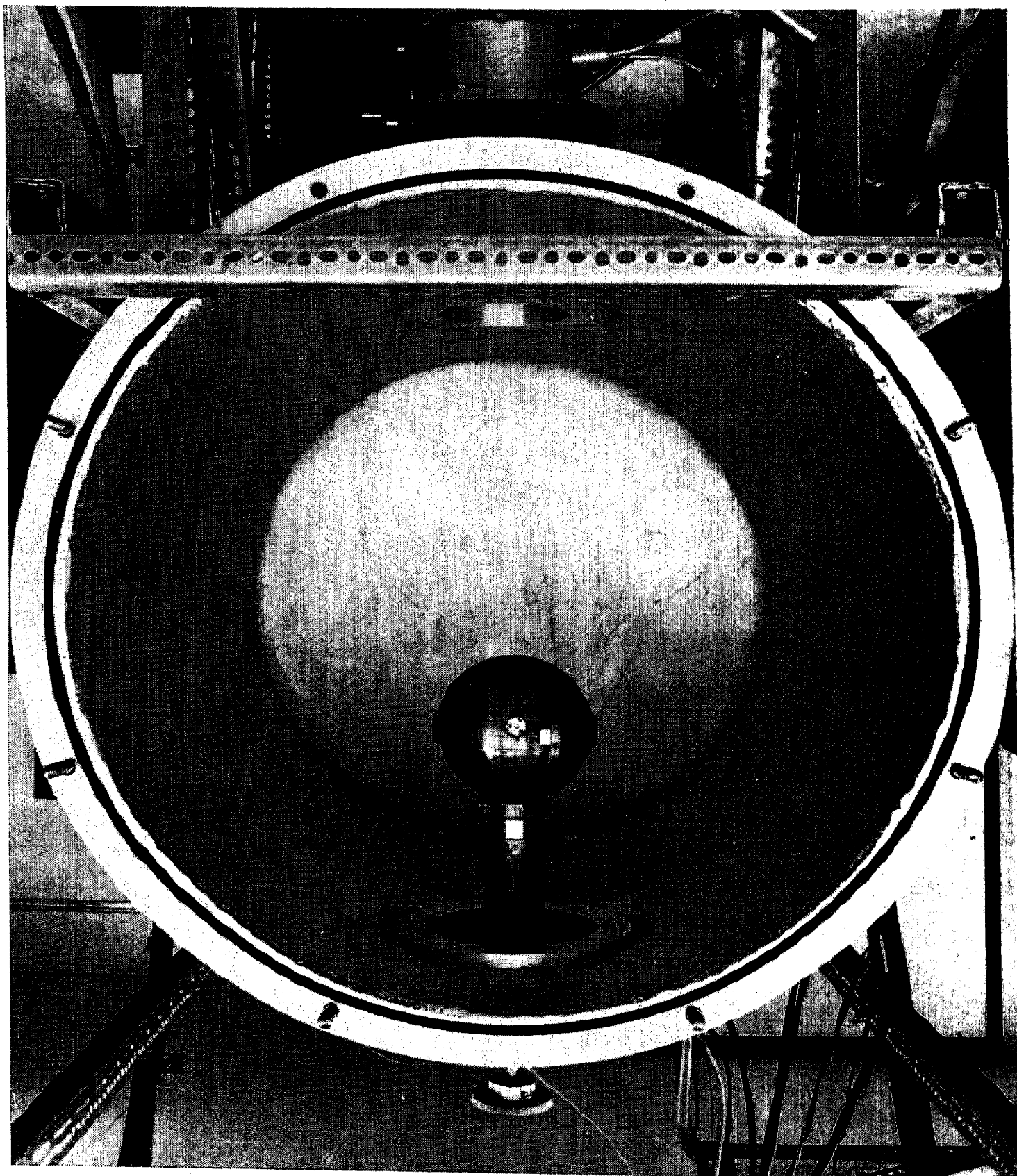


Fig. 7. - Terrella used for experimental determination of threshold rigidities.

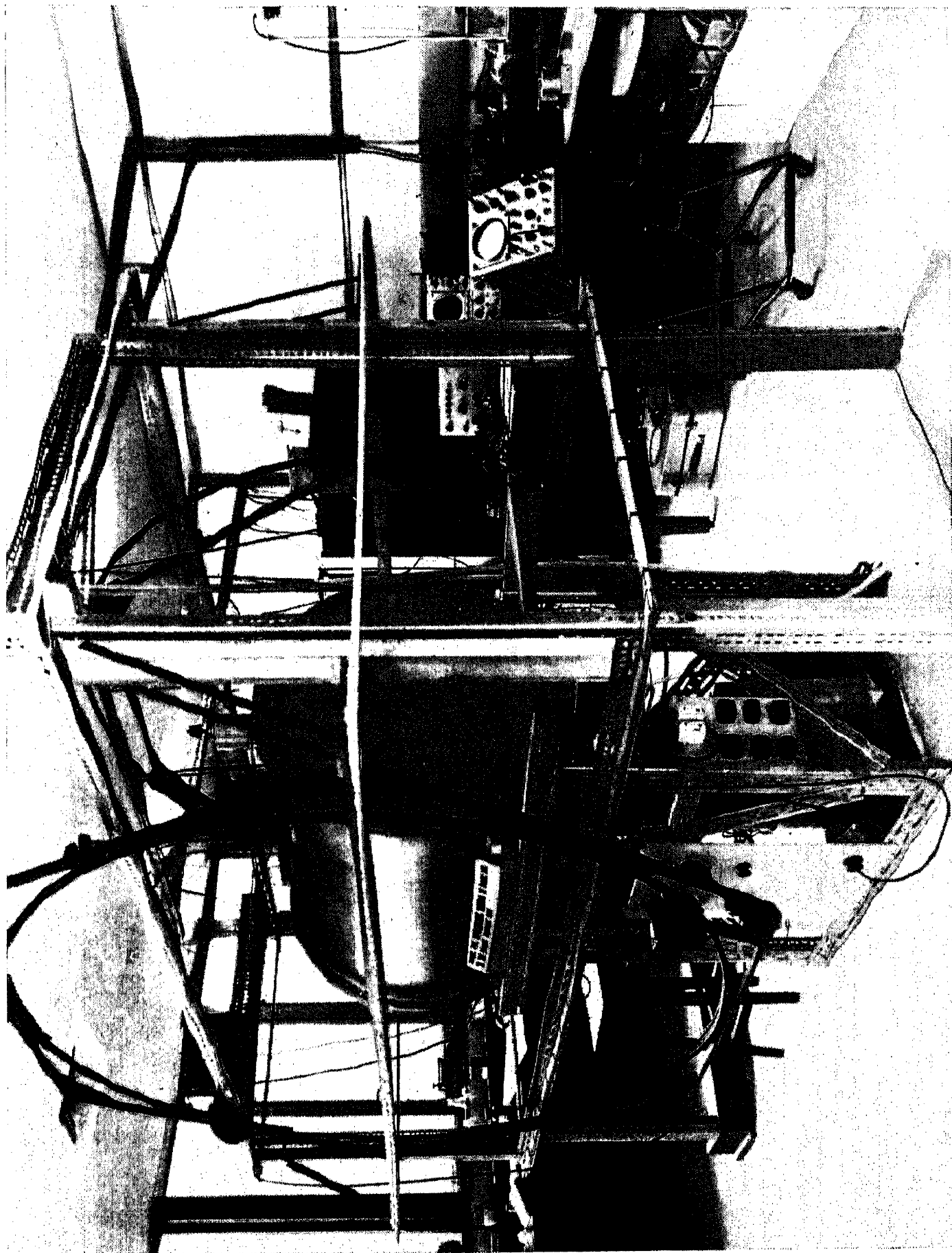


Fig. 8. - Vacuum tank and compensation coils for terrella.

measured as a function of either magnetic field or electron energy depending on which is more convenient. If the electrons have a magnetic rigidity below the threshold they can move only to the terrella. If, on the other hand, their rigidity is greater than the appropriate threshold, they can escape to the tank. Consequently the ratio $i_{\text{terrella}}/i_{\text{tank}}$ changes sharply when the threshold is reached. The corresponding cosmic ray threshold can then be determined from the values of the magnetic moment and electron energy at which the sharp change in the ratio of the two currents occurs.

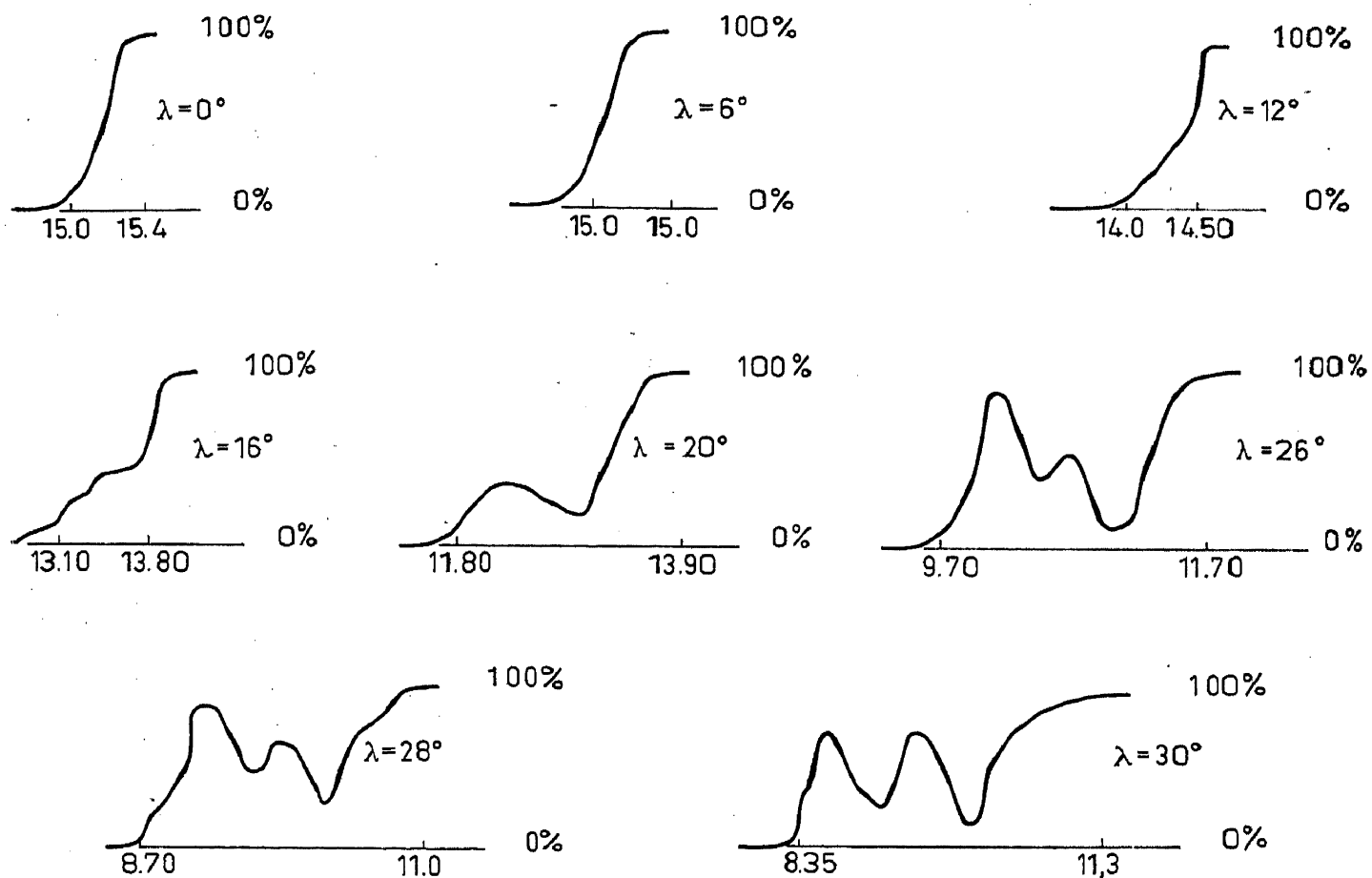


Fig. 9. — Threshold rigidity curves determined for a dipole field, using the terrella.

Figure 9 shows the thresholds measured in this way for several different latitudes in a purely dipole field used as a control. It will be seen that at $\lambda = 0$ the change in current ratio is quite sharp but as we go to higher latitudes the curve becomes more complicated and there is no longer a unique threshold. This effect is due to the penumbra which lies between the Stormer cone, inside which all directions are forbidden, and the main cone, outside which all directions are allowed. The penumbra is made up of both allowed and forbidden directions and this is rejected by the complexity of the curves in Fig. 9.

BLAND is just about to begin measurements of the thresholds using his model of the earth's field incorporating the regional anomalies, and this will be followed by an investigation of the effects of some non-symmetrical external current configurations.

In conclusion, it would be fair to say that we now fully appreciate and understand the effects of the internal parts of the earth's field on the cosmic rays, and, given sufficient computation time, there is no reason why the geomagnetic thresholds should not be fully calculated to any degree of accuracy merited by our knowledge of the field itself. In so far as the external field is concerned, we are only just beginning to appreciate its importance in cosmic ray studies and a proper quantitative appraisal must await detailed information of the current systems involved, but it seems that the changes in threshold rigidity produced by the external field do not in general very much exceed 1 GV, maximum change taking place in the region 50° to 60° latitude.

Solar Protons.

K. W. OGILVIE (*)

Goddard Space-Flight Center

1. – Introduction.

There have been several well known instances of the acceleration by the sun of cosmic ray particles which have been detected at sea level by the traditional methods: the neutron monitor and the meson telescope. It was at one time suggested that this was a very unusual property of solar flares, and indeed that extremely large chromospheric disturbances and unusual conditions near the sun were required to accelerate particles and to allow them to escape. At the present time it is thought that cosmic ray particles, principally protons, are accelerated and escape from the sun much more often than was formerly thought, but that their energy spectrum is such that they can usually be detected only by their effects at high altitudes and high latitudes. Not every class 3+ solar flare is followed by such a « solar beam » interacting with the earth and this situation makes the problem of the solar beam an attractive one, having ramifications which lead far from the traditional areas of cosmic ray research.

2. – Radio observations.

Radiation from the galaxy as a whole must pass through the ionosphere to reach the earth. Since the time scale of galactic changes is longer than that of solar system changes the radio output of the galaxy does not appear to change with time, and therefore observations of the signal strength can show variations which are characteristic of the ionosphere. If a plane wave is incident upon the ionosphere, after traveling a distance d ,

$$E_d = E_0 \exp[-Kd],$$

(*) NASA postdoctoral research associate.

and the absorption in db is

$$A = 20 \log_{10} \frac{E_0}{E_a}.$$

For a single passage of the ionosphere

$$A = 20 \int K dS, \quad K = \frac{2\pi e^2}{mc} \frac{1}{\mu} \frac{N\nu}{\nu^2 + (\omega + \omega_L)^2},$$

where N = electron density, ω = frequency, μ = index of refraction, ω_L = gyro-frequency and ν = collision frequency.

If $\mu = 1$ and $\nu < (\omega \pm \omega_L)$ at high altitude, then $A = (K/\omega^2) \int N\nu dS$. The time constants of the processes are such that equilibrium can be safely assumed at all times. Thus, the absorption in the ionosphere is proportional to $1/\omega^2$ and to $\int N\nu dS$. If the variation of N with height and also the variation of ν with height is known, then the absorption can be calculated. The electron density is here supposed to be due to particles traveling in the earth's field, and incident primarily at high latitudes because of this field. However, the variation of N with height obviously depends upon the energy spectrum, and furthermore the collision frequency as a function of height is not a well known quantity, so the exact height at which the maximum of absorption takes place is still subject to some doubt (*).

An idea of the sensitivity of the absorption can be gained from a rough calculation which shows that the method can detect protons equal to 10 times the normal cosmic ray background incident on the upper hemisphere with energies greater than 100 MeV at 65° geomagnetic latitude. The instrument produced by LITTLE and LEINBACH [3] to measure ionospheric absorption using these principles is called a riometer.

Ionospheric absorption can thus detect the incidence of particles which could never make themselves felt at the ground. The riometer was first made to study the auroral absorption which is such a nuisance to radio communication in the north. It was noticed that a few hours after large solar flares the instrument often shows a characteristic slow increase in absorption, eventually reaching a maximum which subsequently decays over a period of several days. During this time a normal ionosonde, which operates at a much lower frequency and requires a double passage of the ionosphere, is completely « blacked out ». The absorption gradually disappears, and during the latter part of the time, after the sudden commencement of the geomagnetic storm,

(*) A comprehensive calculation on the variation with height of the absorption as a function of latitude and energy spectrum, based on the best values for the collision frequency has been published by BROWN and WEIR [1], and by REID [2].

strong auroral absorption is superimposed. The large value of the absorption observed on some occasions (greater than 15 db at 27 MHz), and the strong latitude-dependence indicate a large flux of low energy particles, an obvious field for study by traditional cosmic ray methods.

Figures 1 and 2, show two records of recent events. The most obvious feature of these is the night-time recovery of signal strenght. This is a local effect caused by the attachment of free electrons. During the day, the incident radiation can detach these electrons, but after sunset at the height of the absorbing layer, about 50 km, they are not available to load the radio waves. The most likely candidate for the attaching molecule is O_2 but NO may play a part [4]. In a calculation using a simple trial spectrum, BAILEY [5] reproduced this day-night effect, predicting a ratio of about 6 to 1 for the absorption (see also [2]).

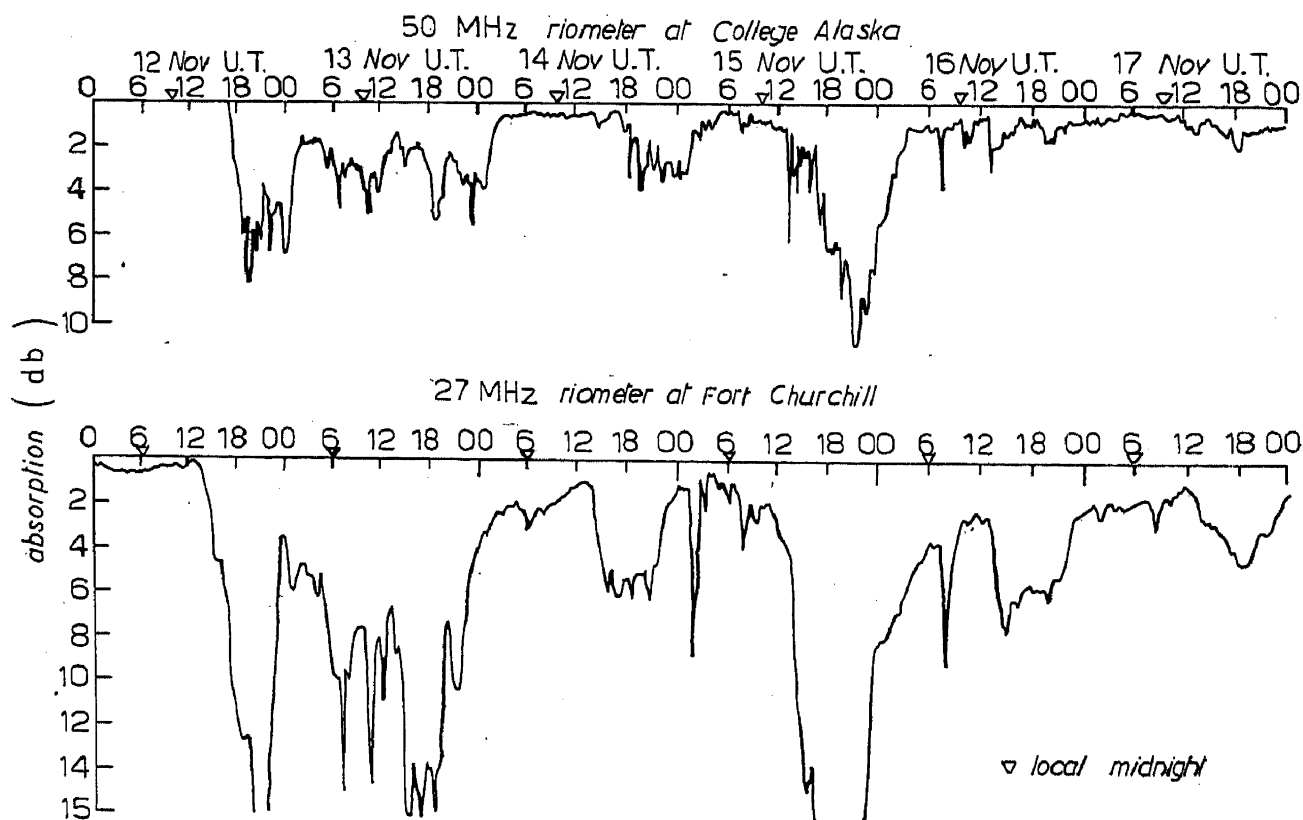


Fig. 1. - Riometer records from College, Alaska, and Fort Churchill, Canada, for the November 1960 cosmic ray event.

In practice, differentiation between polar cap absorption (P.C.A.), as this effect is called, and auroral absorption is not difficult, but in any case propagation by tropospheric scatter is immune to auroral absorption (BAILEY, private communication) as the scattering takes place in the *D* layer.

As was mentioned previously, a normal ionosonde shows blackout conditions during a P.C.A. event. Now the riometer observations do not go back

in time beyond 1957, so that statistics of time variations are poor. These can be improved by using periods of ionosonde blackout, with the further advantage that ionosondes are very well distributed geomagnetically.

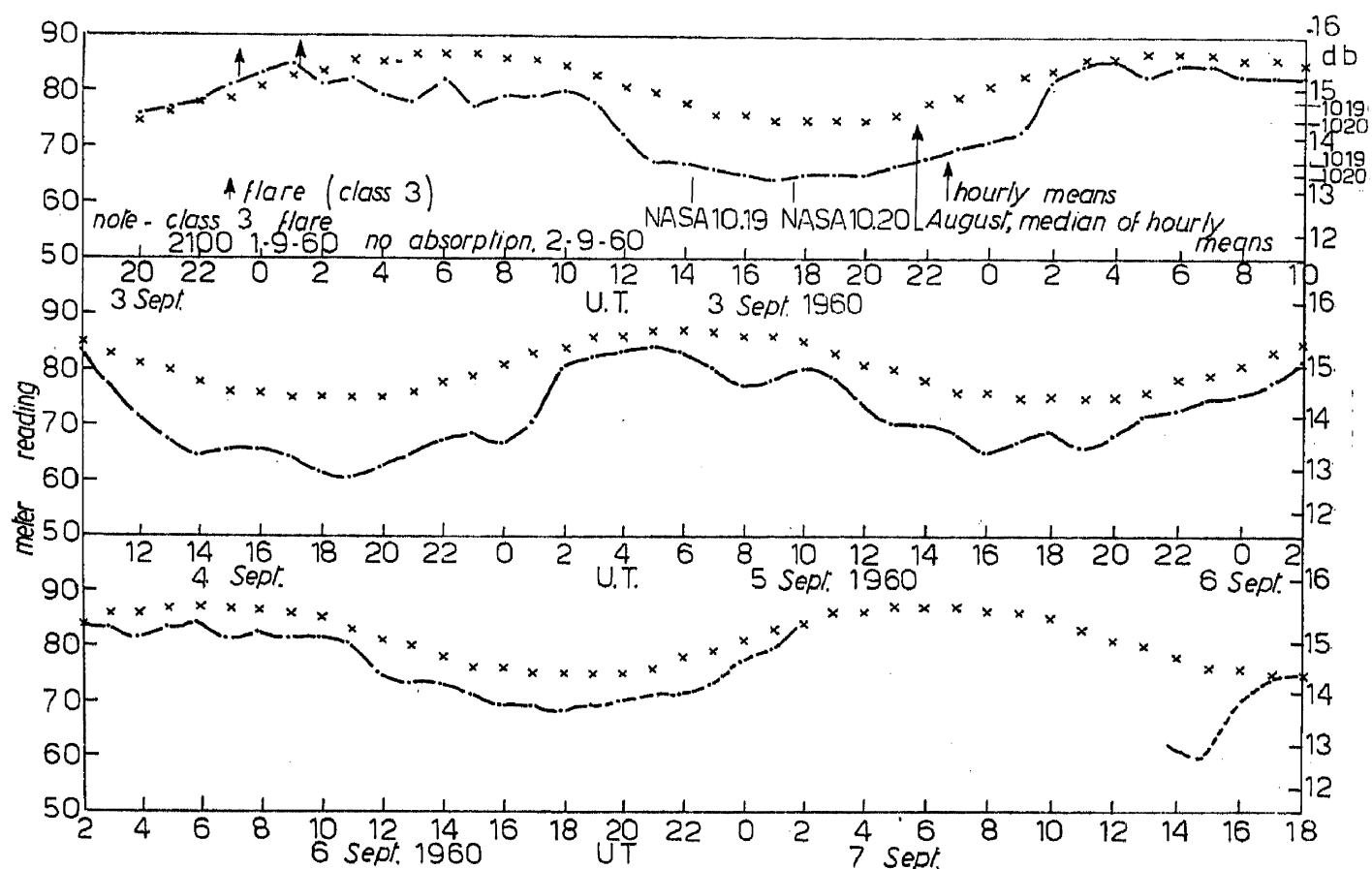


Fig. 2. - Churchill, Canada, riometer record during September 1960 event (27 MHz).

The results of COLLINS, JELLEY and MATTHEWS [6] are an example of the use of this technique. After a suitable comparison with riometer results they set up criteria to diagnose a P.C.A. from ionosonde records. These were:

- 1) Reasonably continuous blackout conditions, that is, no «non»black-out» for more than $\frac{3}{4}$ hour during an event.
- 2) Duration of blackout greater than 3 hours.
- 3) Several stations should see blackout conditions simultaneously.

Using information from the Canadian «net» of stations, they found the time distribution of P.C.A. events from 1949 to 1959 (Fig. 3) and the distribution of events per month during that time (Fig. 4). A satisfactory correlation with the sunspot number is shown, and the tendency for P.C.A.'s to be detected during the northern summer is illustrated.

Statistical examination of P.C.A. events tends to show that the longer the time from solar maximum the greater the effective beam velocity, defined

as the distance from the sun to the earth divided by the time between the flare and the start of the P.C.A. event. Another fact to keep in mind is that five out of the six great cosmic ray events have been detected at a time more than two years away from solar maximum.

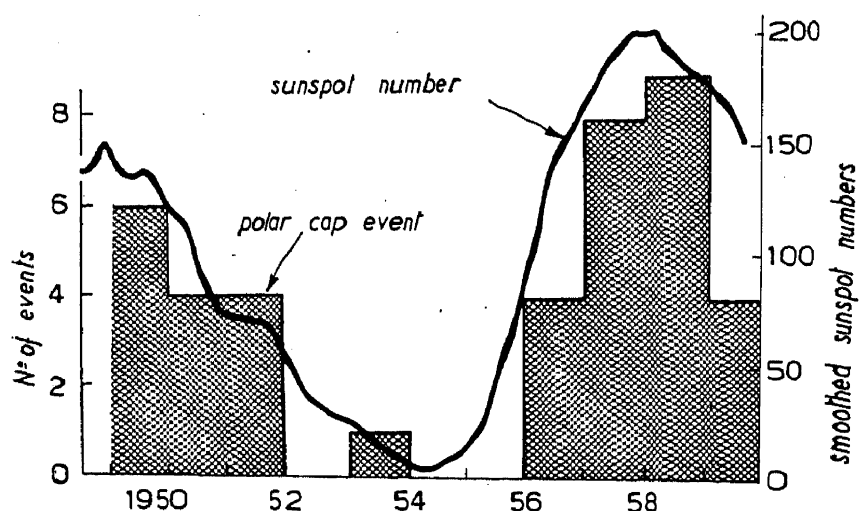


Fig. 3. — Distribution of P.C.A. events from 1949 to 1959. From ref. [6].

An examination of the distribution of P.C.A. events in the two hemispheres using this technique shows a definite time correlation between events in the arctic and the antarctic. It has been found that P.C.A. events in the northern hemisphere are associated with widespread prolonged ionospheric black-outs in the southern hemisphere, presumably also

P.C.A. events. On two occasions (September 26, 1957, and July 29, 1958), out of thirteen studied, no effects could be identified in the antarctic. The northern hemisphere records show low degrees of absorption (2 db and 1 db,

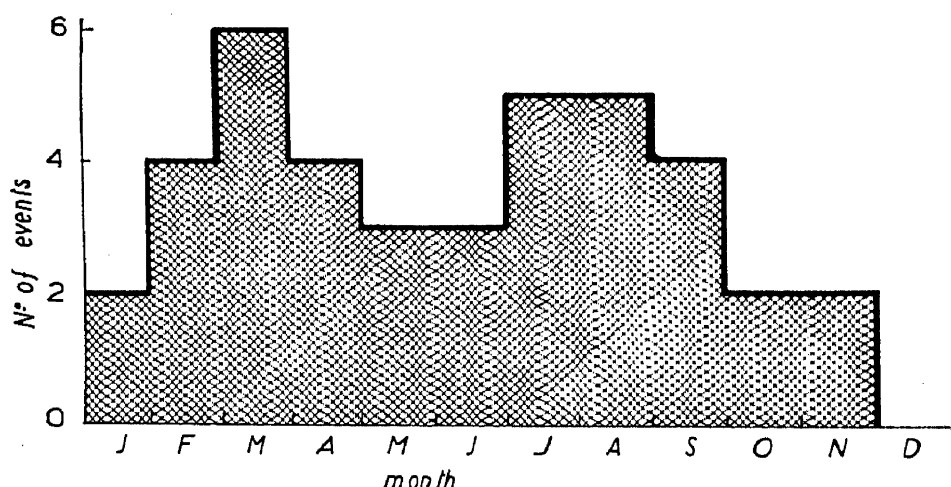


Fig. 4. — Distribution of P.C.A. events by month during the period 1949 to 1959. From ref. [6].

respectively) lasting for about 30 hours each time. These results show a number of remarkable features, but only two of which are discussed here:

- 1) Absorption is frequently found earlier at the geographic pole than at stations of higher magnetic latitude; and
- 2) P.C.A. events are detected during the antarctic summer.

Let us now consider the solar events responsible for these fast particles.

A large chromospheric flare emits a total amount of energy of about 10^{33} erg. Table I gives the approximate division of this energy.

TABLE I. - *Approximate division of the energy from a large chromospheric flare.*

Type	Energy (erg)
Optical emission	$10^{30} \div 10^{32}$
Ultraviolet light	not well known
X rays	$10^{28} < 30 \text{ \AA}$
Radio emission	$10^{24} \div 10^{26}$
Fast particles	$10^{27} \div 10^{30}$
Magnetic storm particles	$10^{30} \div 10^{31}$

The visible light from which the flare takes its name may be so strong that it can be seen directly on the sun's disc. This «white light» emission is thought to be synchrotron radiation from fast electrons. The counterpart of this in the radio region is synchrotron radiation, continuum emission over a wide frequency range, which is found to be strongly correlated with the P.C.A. events. Since the intensity of this «type IV» continuum varies rapidly with frequency, the best way to observe it is to record the emission over as wide a frequency range as possible. Considering the 31 P.C.A. events between February, 1956, and July, 1959, KUNDU and HADDOCK [7] have shown that of 28 P.C.A. events which occurred during periods of centimetric observations all 28 were associated with centimetric wave outbursts (Table II). This table shows that 83 percent of intense broad-band centimetric outbursts are associated with P.C.A. events. After a type IV event a center of emission low down in the chromosphere can be formed, lasting for several days. It is also evident that type IV emission detected at both centimetric and at metric wavelengths can occur without P.C.A. events being detected at the earth (*).

In view of the interest now being taken in shock waves in plasmas, it is interesting that type III bursts, which are considered to be caused by the ejection of particles from the sun with velocities of the order of $c/10$, are associated with flares. This correlation increases with the magnitude of the flare. If these emissions are shock waves, they may be of interest for the interpretation of other cosmic ray phenomena, for example, Forbush decreases. Type III radio bursts are not correlated with neutron monitor increases, as is shown by the present author and by McLEAN (private communication (**).

(*) DENISE (ref. [8]) gives 30 percent as the proportion of 3 and 3+ flares associated with type IV emission.

(**) See also *Proceedings of the International Conference on Cosmic Rays and the Earth Storm*, Kyoto, Japan (September 1961); *Suppl. Journ. Phys. Soc. Japan*, 17, (1962).

In almost every one of the 45 instances since 1956 when there has been a P.C.A. event, a flare on the visible disc of the sun can be considered to be the source. Examinations of these flares in the broadest sense of the term have

TABLE II (*). - *Association of radio outbursts with polar cap absorption events*
(May, 1959-July, 1959).

Radio outbursts	Centimeter-wave intensity	Number of radio events	Polar cap absorption events		Percentage of polar cap absorption association
			Yes	No	
Broad-band centimeter-wave outbursts independent of meter-wave outbursts occurrence	Intense and moderate	34	12	2	83
	Intense only	12	10	2	83
Type IV meter-wave outbursts independent of centimeter-wave outbursts occurrence		19	10	9	53
Type IV meter-wave outbursts associated with broad-band centimeter-wave outbursts	Moderate	7	0	7	0
	Intense	12	12	2	83

(*) From ref. [7].

shown very little that can be used to identify which ones emit particles. The particles are probably accelerated by a Fermi process in the region of high turbulence close to the site of the optical flare. The region emitting the radio waves can be shown to move away to a considerable distance (of the order of 1 solar radius) with a velocity of the order of 1000 km/s [8], and this is consistent with the required lifetime of the particles to provide the correct duration for the radio continuum emissions. Flares which emit particles reaching the earth appear to occur preferentially on the western side of the sun's disc, but statistics are poor. A preponderance of events due to flares on the western side might be expected if the particles are thought of as following lines of force drawn out and bent by solar rotation. However, this view is by no means accepted generally. For example, the solid angle into which a solar beam is projected is obviously not small, but nevertheless there are very few events (one or two at the most), which could be attributed to flares taking place beyond the western limb. If there is sometimes a disordered magnetic field region at a radius of several A.U., as was introduced by SIMPSON and PARKER [9] to explain the decrease of cosmic ray intensity following the February 23, 1956, event,

it does not seem to be a good reflector all the time. Flares near the center of the sun's disc are known to be statistically associated with the largest magnetic storms. OBA YASHI and HAKURA [10] give the data shown in Table III.

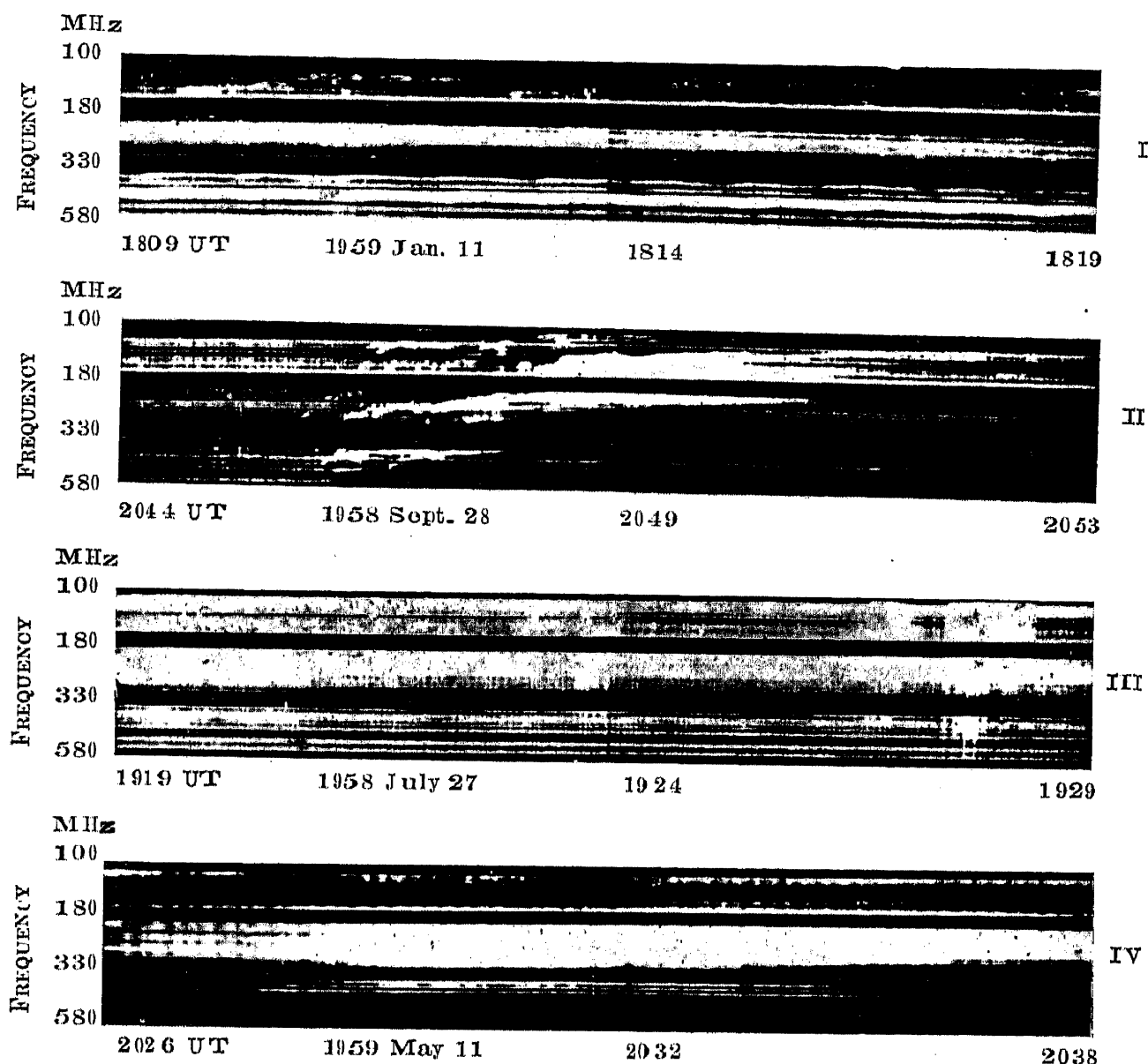


Fig. 5. - The four spectral types of solar radio bursts. (I) noise-storm bursts; (II) slow-drift bursts; (III) fast-drift bursts; (IV) continuum bursts. From G. SWARUP, P. H. STONE and A. MAXWELL: *The association of solar radio bursts with flares and prominences*, in *The Astrophysical Journal*, 131, 725 (1960).

As in all statistical studies, due attention should be paid to the criteria of selection in examining the conclusions. This seems especially to be true of solar flare observations, as may be appreciated by referring to Dodson's survey [11] of solar flares during the International Geophysical Year, and to the paper of Warwick and Haurwitz, reference [11b].

In view of the importance of solar protons to manned space flights it would be very convenient if it were possible to predict when an event would take

place. Of course, for long space flights this would not be of much use, the only final method of protection being shielding; but for orbital flights, prediction could possibly be used to save weight. ANDERSON [12] has made a study of this problem. He finds that the large flares that produce protons

TABLE III (*). — *Variation of time interval between flare and arrival of protons and magnetic storms as a function of magnetic position.*

Position on sun's disc	East $90^\circ \div 30^\circ$	Center	West $30^\circ \div 90^\circ$
Average (Protons)	18.3 h (**)	7.5 h	3.6 h
(Magnetic storm)	37	38	37
ΔH	139 γ	345 γ	106 γ

(*) From ref. [10].
(**) See also ref. [11b].

occur in long-lived and complex sunspot groups. The number of sunspots in a group is of no use in prediction, but there does seem to be a correlation between penumbral area in spot groups and the occurrence of flares producing solar protons (Fig. 6, 7, 8). This large penumbral area is said to be an unusual feature of spot groups. No attempt has been made by ANDERSON or anyone else to provide a theoretical basis for this correlation, if it is supposed real.

Because solar proton events last for several days, the commencement of the magnetic storm often occurs during the time that the particles are still coming to the earth. This is a fortunate circumstance because time and latitude variations of the intensity allow us to use the protons to study changes in the field.

The classical geomagnetic storm has a sudden commencement (SC) at which the field H begins to increase. After a period of about one hour, the field undergoes a decrease, lasting for several hours, after which it slowly returns to normal.

In an extensive survey of this subject BACHELET *et al.* [13] have found that *Forbush decreases* are normally associated with magnetic storms of the classical type, or ones in which there is no main phase. Magnetic storms with no sudden commencement, or in which the SC is merely followed by irregular variations, are not associated with Forbush decreases. These magnetic storms not associated with Forbush decreases are typical of solar minimum, and may be correlated with so called M regions, not with typical activity centers. If this is the case, emission of the solar protons from activity centers should be associated with normal magnetic storms, and this is found to be generally correct.

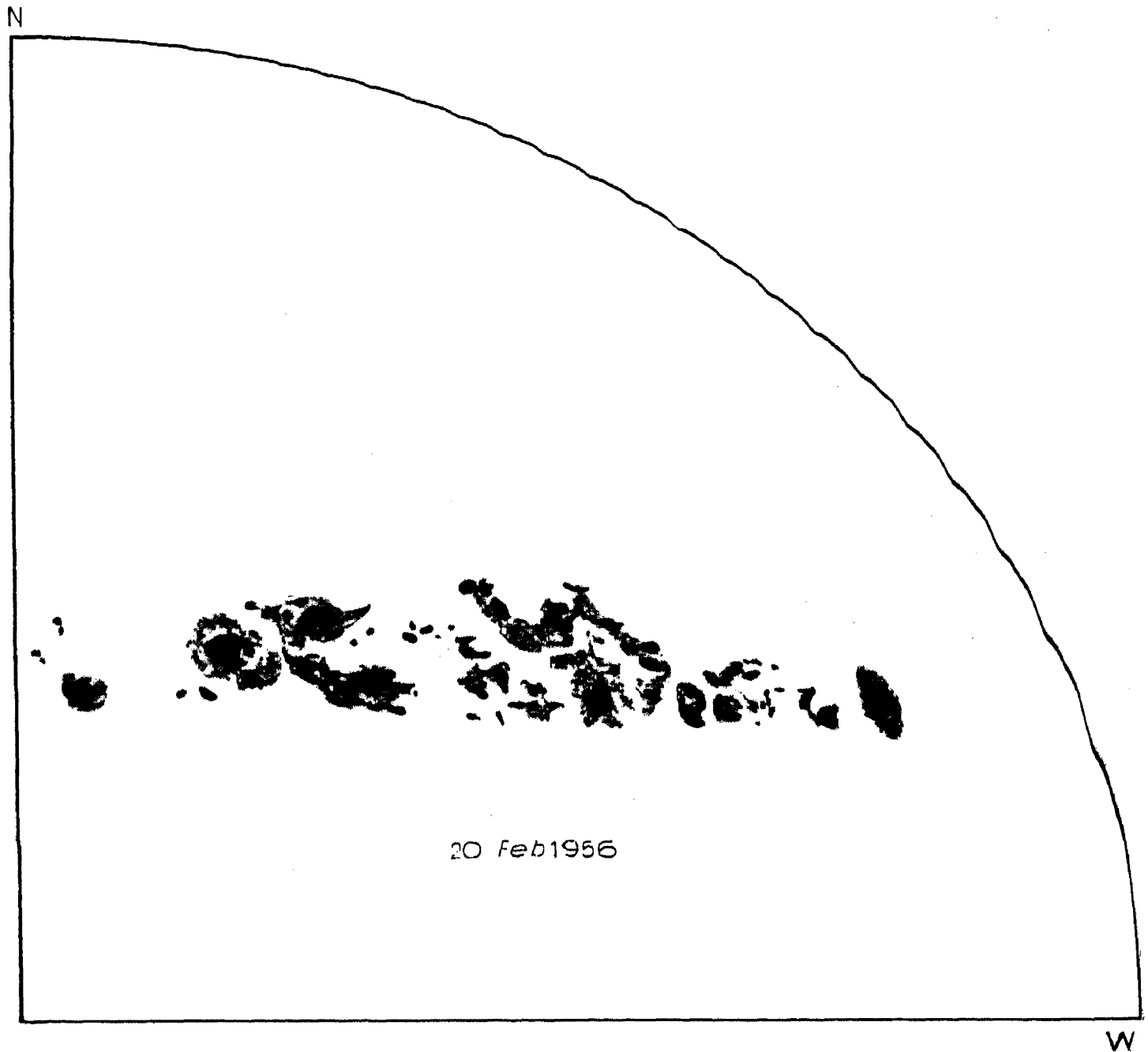


Fig. 6. — The sun on February 20, 1956, from a photograph in *Publication of the Astronomical Society of the Pacific*, 68, 273 (1956). The figure caption states: «The long stream of spots ... contained five separate groups as indicated by their magnetic fields. The intense flare observed at Tokyo on February 23 was near the center of this stream which was then very close to the west limb.» The large penumbral areas are very prominent. From ref. [12].

The rise in field during the initial phase of the storm is widely held to be due to the compression of the earth's field by an incoming front. ODAYASHI [14, 14b] has made a calculation on this subject. Noting that the southward edge of the polar blackout moves toward the equator during the main (negative) phase of the storm, he carries through a Størmer calculation using a modified mag-

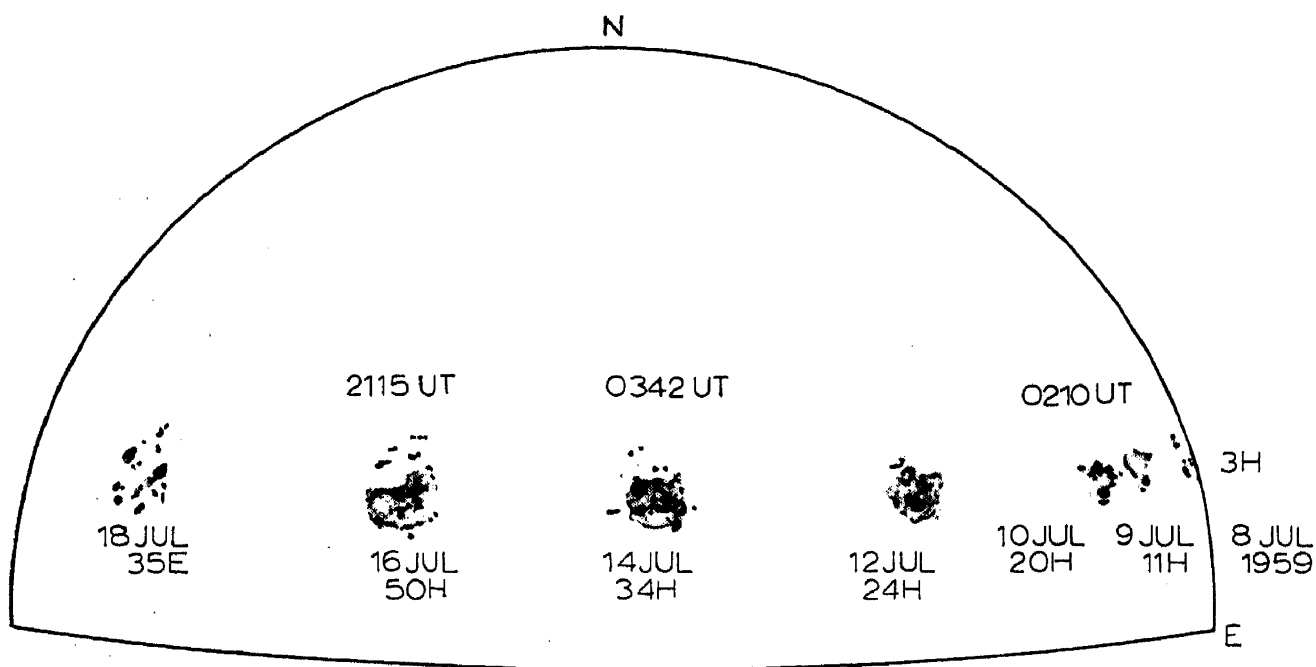


Fig. 7. - Two days before the first great July, 1959, solar proton emission, the parent spot group appeared at the limb. It contained only three spots when first seen at Athens, but the penumbral development was already very great. From ref. [12].

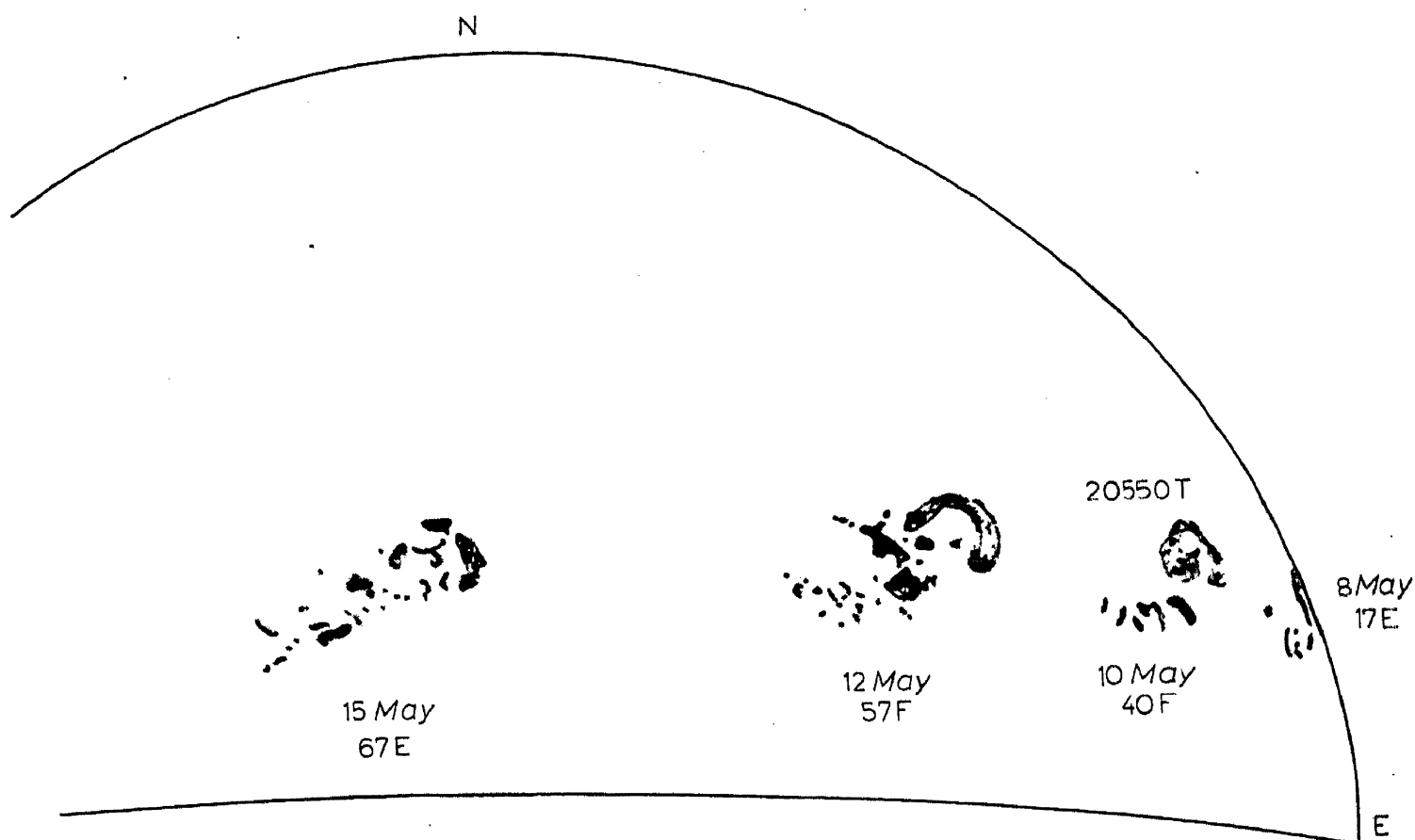


Fig. 8. - The flare of May 10, 1959, gave rise to an intense flux of solar protons that reached the earth. Two days before the flare, the large penumbral areas of the sunspot group could be clearly seen, greatly foreshortened, near the east limb. From ref. [12].

netic potential. Instead of simply adding a constant field

$$P = M \left(\frac{1}{R^2} + AR \right) \sin \varphi ,$$

where A is a constant and P is the potential, he superimposes the following two conditions successively:

$$H_r = 0 \text{ at the boundary of a cavity of radius } R_0 ,$$

$$H = 0 \text{ at the boundary of a cavity of radius } R_0 .$$

This gives two potentials

$$P_1 = M \left(\frac{1}{R^2} + \frac{2R}{R_0^3} \right) \sin \varphi ,$$

and

$$P_2 = M \left(\frac{1}{R^2} - \frac{R}{R_0^3} \right) \sin \varphi .$$

In the first case (P_1), no line of force crosses the boundary of the cavity $R = R_0$; and in the second (P_2), the field is drawn out so that the lines of force cross this boundary at right angles. The particles outside the cavity may enter if they have sufficient energy, the distortion of the earth's dipole field by the geomagnetic storm being such as to increase it at the equator during the positive phase, and reduce it during the main phase:

$$H_1 = \frac{2M}{R_0^3} \cos \varphi ,$$

$$H_2 = - \frac{M}{R_0^3} \cos \varphi .$$

The radius of the cavity is thus the parameter which, on this theory, fixes the field change. The first integral of the equations of motion of the proton in a field of this potential is used to find Størmer allowed and forbidden regions. The results of the calculation of the cut-off for vertical incidence as a function of latitude using R_0/R_s as a parameter, are shown in Fig. 9 and 10.

This problem is also discussed by ROTHWELL [15]. In her treatment, the solar gas cloud enters the earth's field and approaches to a distance of several earth radii (xR_s) where particles are supposed to leave it and travel along the lines of force of the earth's dipole field. The dipole field is supposed to be unaffected by the cloud at distances less than $R = xR_s$. Since there is no

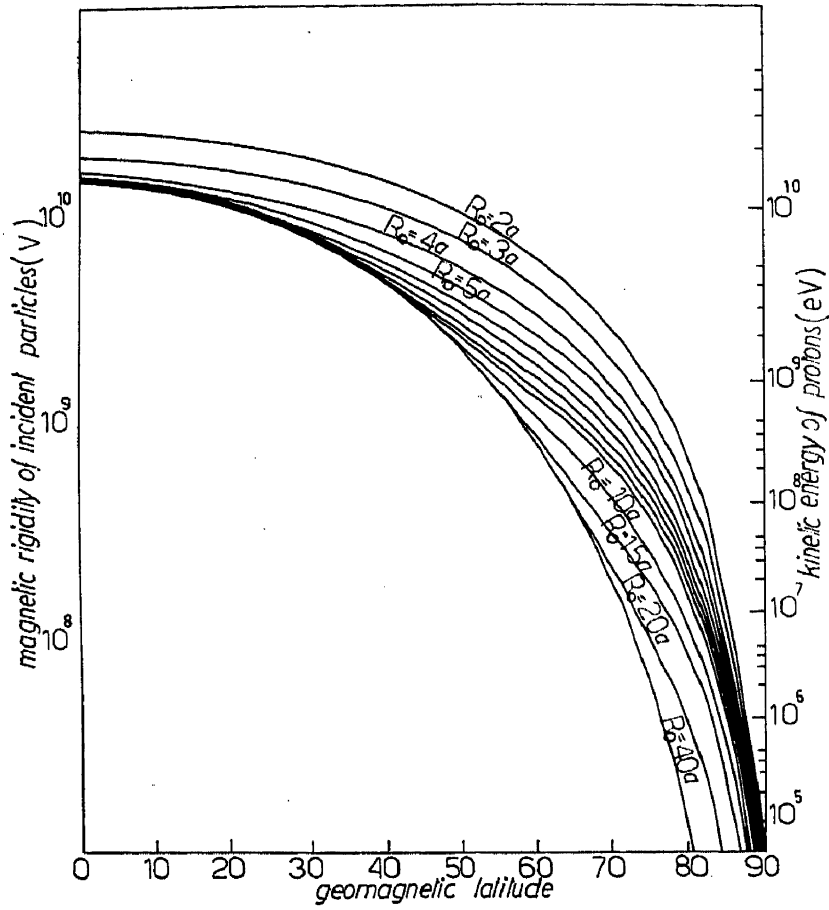


Fig. 9. - Geomagnetic cut-off rigidity of particles P_1 vs. geomagnetic latitude for various R_0 . From ref. [14]. Geomagnetic potential

$$\Omega = M \left(\frac{1}{R^2} + \frac{2R}{R_0^3} \right) \sin \varphi,$$

$$\cos^2 \varphi_0 = \frac{2\alpha\gamma}{(eM/mv)^{\frac{1}{2}}(1 - \alpha^3/R_0^3)}.$$

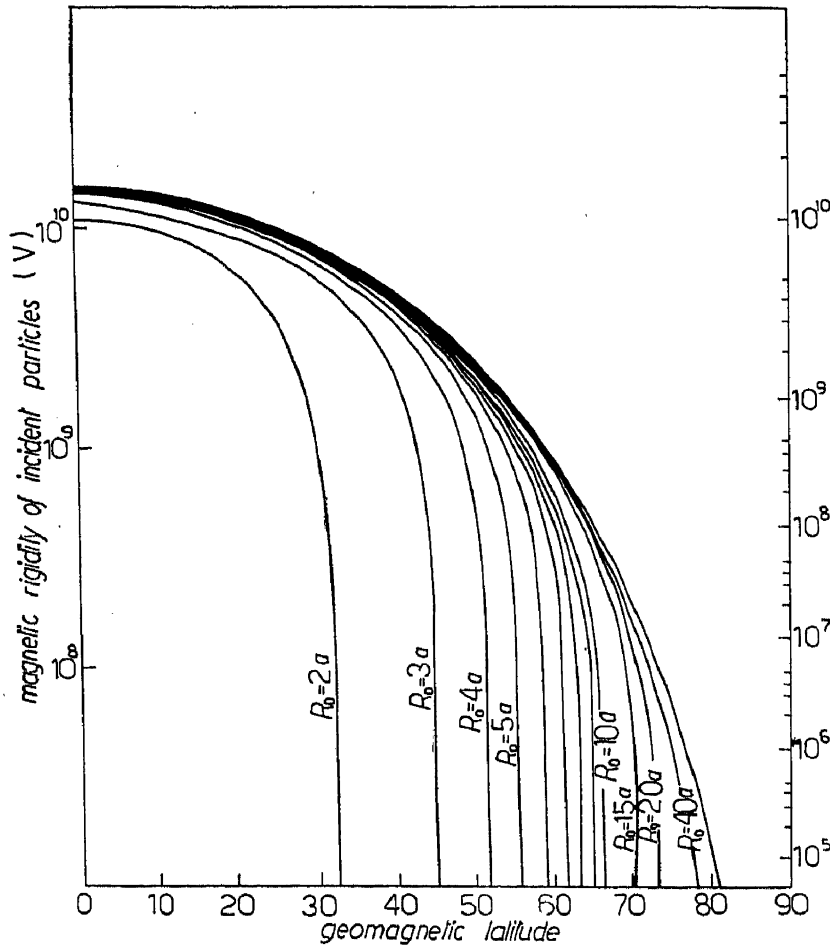


Fig. 10. - Geomagnetic cut-off rigidity of particles P_2 vs. geomagnetic latitude for various R_0 . From ref. [14]. Geomagnetic potential

$$\Omega = M \left(\frac{1}{R^2} - \frac{R}{R_0^3} \right) \sin \varphi,$$

$$\cos^2 \varphi_0 = \frac{2\alpha\gamma}{(eM/mv)^{\frac{1}{2}}(1 + \alpha^3/2R_0^3)}.$$

« compression » in this model, the cut-offs are only reduced, and the positive part of the storm is not explained. Essentially, the function of the cloud is to bring particles to certain distances and let them start from a point where the potential is higher than at infinity. This idea separates entirely the dynamic and magnetic properties of the cloud, and is thus not very realistic. On the other hand, Obayashi's treatment reproduces *semi-quantitatively* the main features which are actually observed by introducing an *a priori* assumption.

These matters have been discussed before the main topic because the situation during a solar proton event is very complex. It is essential to bear in mind as many aspects of the geophysical situation as possible, the magnetic field disturbances in particular being essential clues to the processes occurring.

3. — Methods of observing solar protons.

To provide an idea of the relative sensitivities of detectors, some figures calculated by WEBBER (private communication) are given in Table IV. This very illuminating tabulation shows why the protons incident during P.C.A. events were discovered by the riometer. It is the most sensitive sea level detector of incident charged particles, and this is especially true when there

TABLE IV. — *Relative sensitivities of detectors.*

Instrument	Altitude and position	Assumed minimum detectability	Increase $\times$ C. R. Greater than 100 MeV	
			P^{-6}	P^{-8} (*)
Ion chamber	$\lambda = 50^\circ$, S.L.	$+ 2 \pm$	10 000	50 000
Neutron monitor	$\lambda = 60^\circ$, S.L.	$+ 2\%$	250	1 000
Single counter	$\lambda = 65^\circ$, 100 g/cm ²	$+ 5\%$	25	100
Ion chamber	$\lambda = 65^\circ$, 100 g/cm ²	$+ 5\%$	15	70
Riometer	$\lambda = 65^\circ$, S.L.	$\sim \frac{1}{2}$ db	12	12
Riometer	$\lambda = 67.5^\circ$, S.L.	$\sim \frac{1}{2}$ db	2	0.5
Riometer	$\beta = 70^\circ$, S.L.	$\sim \frac{1}{2}$ db	0.2	0.02
Single counter	$\lambda = 65^\circ$, 10 g/cm ²	$+ 5\%$	0.2	0.5
Ion chamber	10 g/cm ²	$+ 5\%$	0.08	0.2
Satellite or rocket	2 g/cm ²	$+ 10\%$	0.01	0.004

(*) The two columns are obtained by using two spectra for the incident particles, namely power law rigidity spectra with exponents -6 and -8 .

is a steep spectrum. Thus, to get more detailed information about the incident particles—and the riometer can, in principle, give a good deal of information

about the spectrum and intensity—high altitude observations are essential. The available methods of reaching high altitudes are given in Table V.

TABLE V. — *Characteristics of methods of observing at high altitudes.*

Method	Maximum altitude (g/cm ²)	Duration	Power
Aircraft	~ 50 g	~ 10 h	large
Rubber balloons	~ 10 g	up and down	small
Plastic balloons	~ 4 g	~ 10 h	small
Rockets	zero	~ 2 min	small
Satellites	zero	long	small
Space probes	zero	long	small

Aircraft were useful when a large power source was often required for complicated experiments, but the use of the transistor and miniature components has eliminated the need for them. Rubber balloons are superseded because they do not float at high altitude, and in fact most of the work being done in this field uses two methods; constant volume plastic balloons and rockets.

It may seem surprising that the rocket, with its inherent limitation of short flight, is of any use compared with the satellite. However, the available

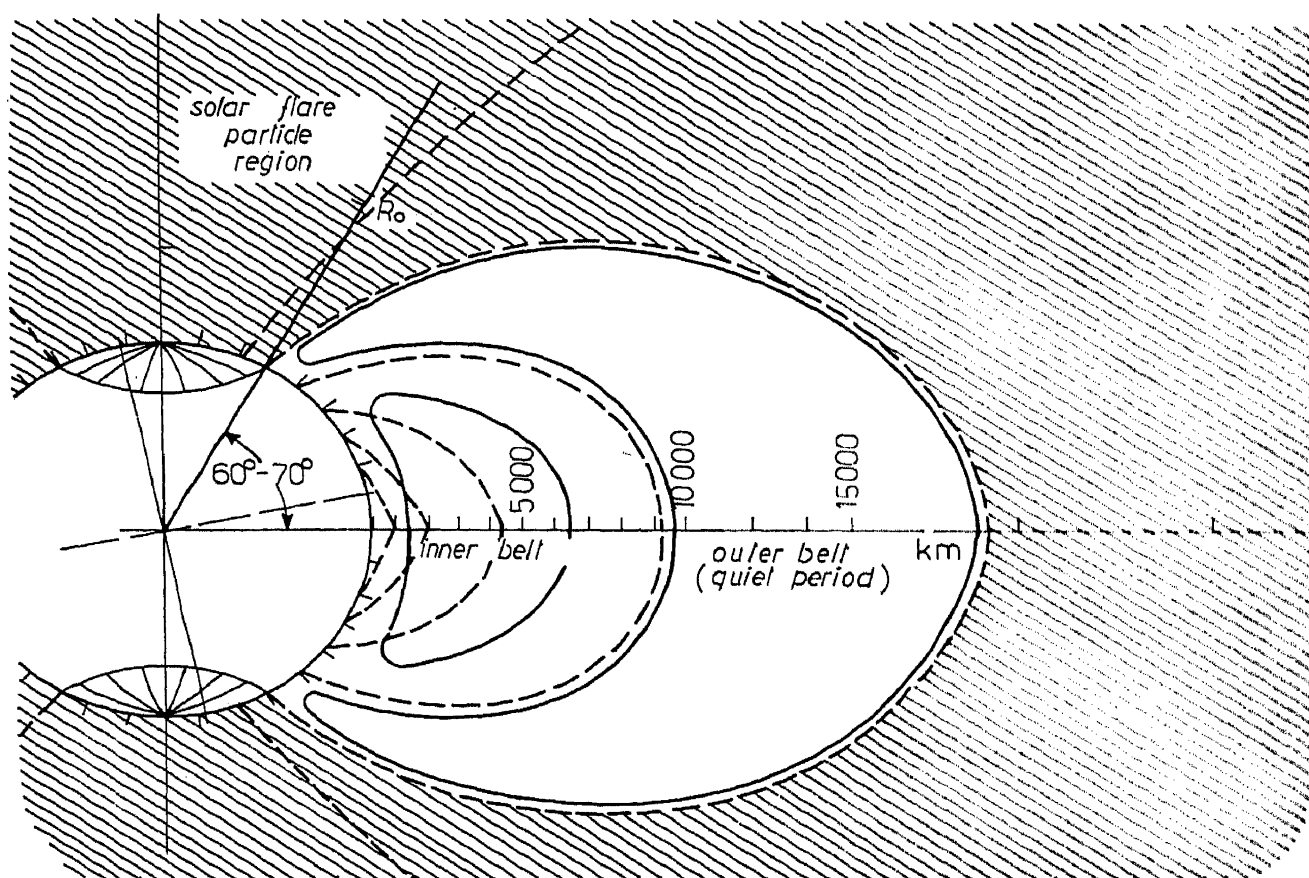


Fig. 11. — Geometry of radiation belts and region to which solar flare particles penetrate.

satellites are often in orbits which are not inclined at a large angle to the equator. They do not precess in inclination, so the situation is that unless a satellite orbit is inclined at a large angle to the equator, or of great eccentricity such as Explorer XII, it cannot sample the volume of space through which particles approach the polar regions (Fig. 11).

4. - Detectors.

What should be borne in mind in designing an experiment to detect solar protons? Figure 12 shows the penetration depth of protons in the earth's atmosphere, the corresponding dipole cut-off latitude, and dE/dx . For a rocket experiment, the minimum energy which is to be considered is set by the detector; but for a balloon, the atmospheric cut-off will normally prevent meas-

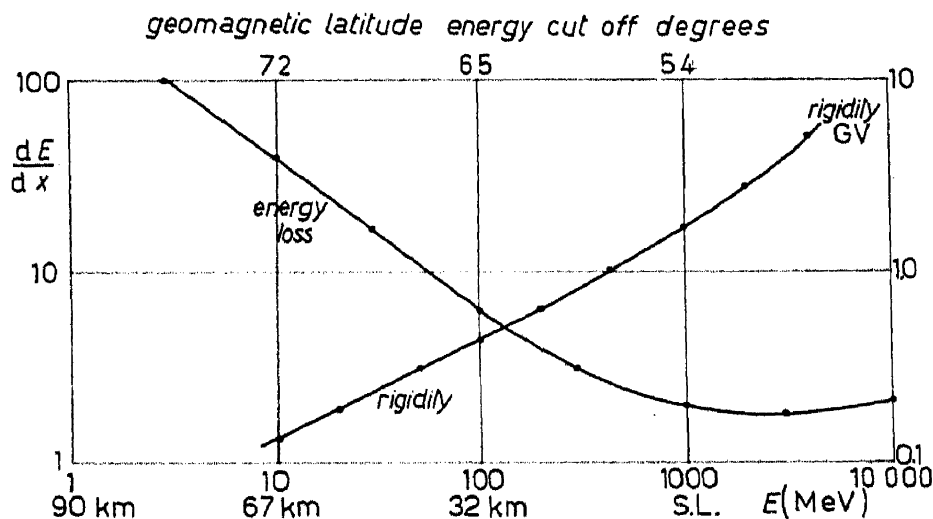


Fig. 12. - Energy loss and rigidity as a function of energy. Depths of penetration into the atmosphere and magnetic co-ordinates are shown.

urements below about 100 MeV. Even an isotropic angular distribution of protons in the upper hemisphere above the atmosphere rapidly becomes peaked at lower altitudes. Thus, at 10 g/cm², say, most of the protons to be detected will have an ionization of several times minimum and will be travelling vertically. This means that the combination of calibrated ion chambers and a Geiger counter or counter telescope can differentiate between protons and other particles where dE/dx is greater than minimum. This arrangement has been extensively used by WINCKLER [16], and by ANDERSON [17]. See Table VI and Fig. 13.

For example, during the period August 22 to August 25, 1958, balloon flights were carried out at Churchill, Canada, Fairbanks, Alaska, and Min-

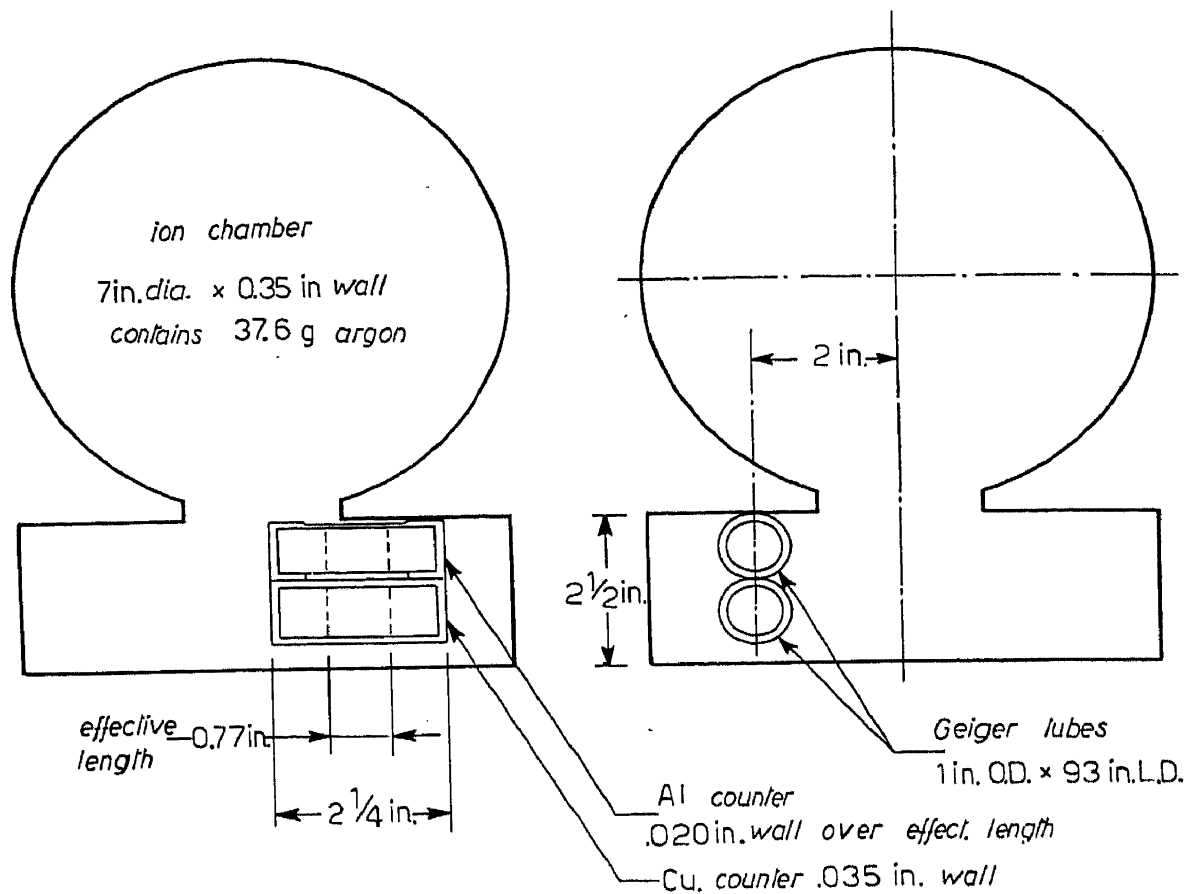


Fig. 13. - Geiger counter and ionization chamber used with balloons by Winkler.
From ref. [16].

neapolis. The spectrum can be obtained from the plot of the absorption curve of the radiation during the balloon ascent. ANDERSON *et al.* [17] assume:

- 1) There are only protons present.
- 2) They are isotropic in the upper hemisphere.
- 3) They are absorbed by ionization loss alone.
- 4) There is no scattering.

By introducing a differential energy spectrum of the form $(K/E^n)dE$, and a range-energy relation $\text{Range} = E^u$, the absorption curve can be calculated as a function of the exponent of the spectrum, and a fit obtained. The result in the particular case was $n = 5$, a typical value, and this applied from 250 MeV corresponding to the depth at which the counting rate started to rise, to about 100 MeV, the air cut-off at high altitude. The agreement of the calculated absorption curve with experiment did not by itself confirm the nature of the radiation as protons, but electrons could not have produced the high value of dE/dx which is inferred from the ratio of the rate of discharge of the ion chamber to the rate of the single counter, about three times that given by

normal cosmic rays. This is confirmed by the high degree of directional correlation observed by means of the counter telescope.

TABLE VI. — *Critical data on the monitoring detector unit.*

Detector	Geometry and dimension	Solid angle factor	Wall thickness and material	Absorption to the vertical	Notes
Ion Chamber (*)	sphere: 17.75 cm dia.	Omni-directional projected area 250 cm ²	Aluminum 0.089 cm	0.4 g/cm	Electrometer charge per normalized pulse $0.67 \cdot 10^{-10}$ coulomb
Copper counter	cylinder: 2.36 cm dia. 1.96 cm effective length	Omni-directional projected area 5.8 cm ²	Copper 0.089 cm	1.5 g/cm	Recovery time 150 μ s Cosmic ray efficiency > 94%
Aluminum counter	cylinder: 2.36 cm dia. 1.96 cm effective length	Omni-directional projected area 5.8 cm ²	Aluminum 0.051 cm	1 lg/cm	Recovery time 150 μ s Cosmic ray efficiency > 94%
Telescope		Solid angle factors 2π omni-directional intensity (5.3 ± 0.5) cm ² sr 2 cos θ intensity (3.6 ± 0.4) cm ² sr		1.5 g/cm	Resolving time of coincidence circuit $\sim 3 \mu$ s

(*) From ref. [16].

(**) Ion pairs/cm² s at NTP air = $12.3 \cdot 10^3$ ion pulses/s.

Although this method shows that the observations are consistent with a beam of 100 % protons, it cannot rule out the presence of mixture of high energy electrons and α particles, particularly since adding electrons reduces dE/dx and adding α -particles increases it. It would be possible to account for the observations entirely by a mixture of high energy electrons and α -particles but this can be shown to be unlikely in a simple way. The transit times would be so different that the ratio observed between the detector rates would vary during the beginning of an event, which is not observed.

A good confirmation of the nature of the particles is provided by simultaneous observations at different points. As an example of this technique

consider the same event, as shown in Fig. 14. The flights at Churchill, Minneapolis, and Fairbanks focus attention on the following points.

- 1) The long period of observation at constant altitude provided by a successful flight.
- 2) The absence of an increase at Minneapolis.
- 3) The Forbush decrease observed at high altitude.

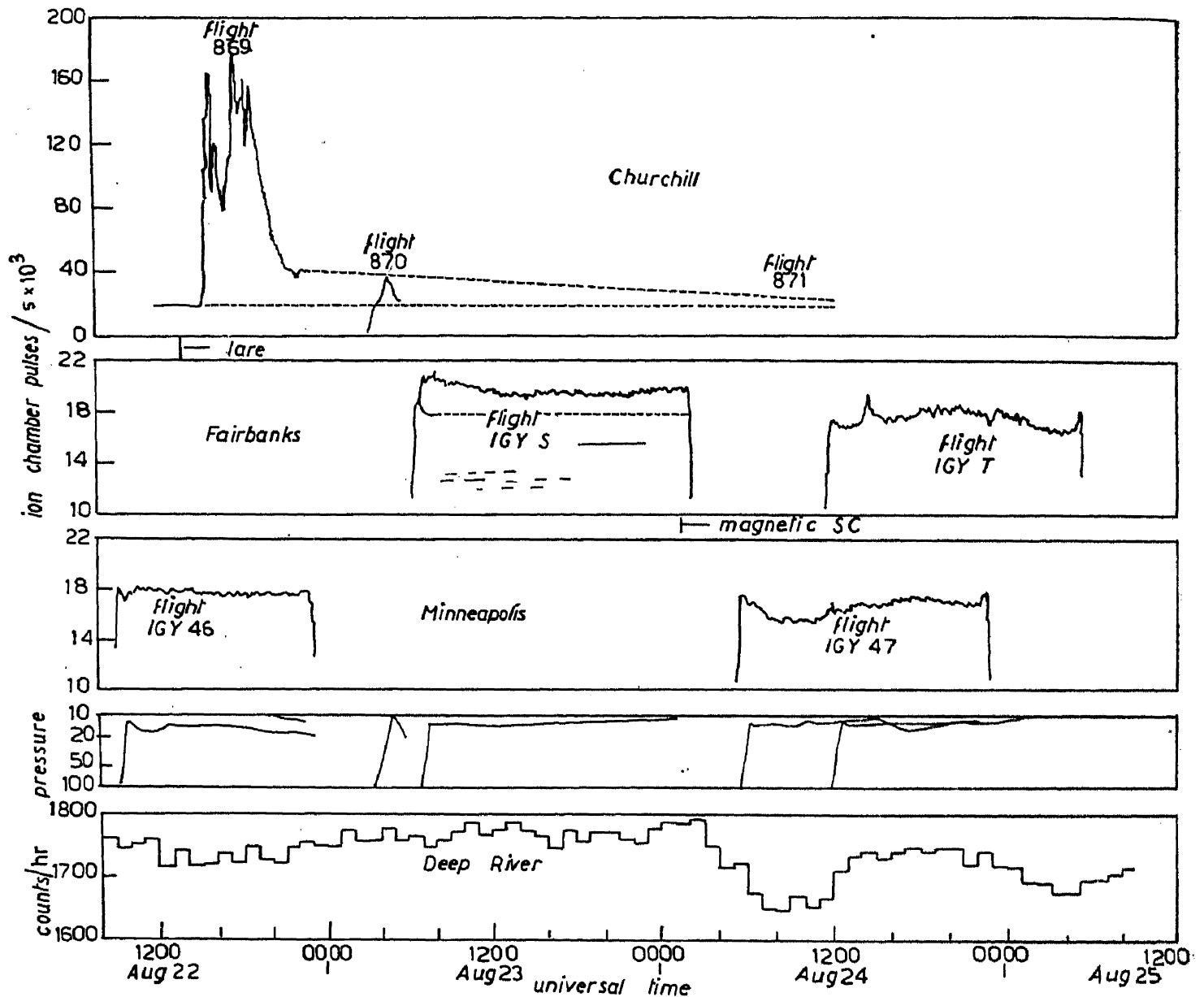


Fig. 14. — Balloon observations and correlated data during the cosmic-ray event of August 22, 1958. Top line, observations at Churchill; next lower, observations at Fairbanks; center, observations at Minneapolis; next lower, balloon pressure curves; and bottom, Deep River sea-level neutron monitor. From ref. [15].

The absence of the increase at Minneapolis shows that there were no protons incident there with enough energy to overcome the cut-off. In addition, this also puts an upper limit to the abundance of radiations not subject to a

cut-off, namely neutrons and photons. For this particular event these limits were 12% with energy greater than 40 MeV and 10% with energy greater than 500 keV, respectively. Again the event seems to be typical, because subsequent measurements have confirmed that the composition of the beam is always principally protons.

Observations at high altitude and medium latitudes—for example, those at Minneapolis—have a great additional bonus. If particles do appear there, and are identified as protons, then this becomes evidence of a cut-off change, Fig. 15 and 16. In the events of August 22, 1958, there was no geomagnetic

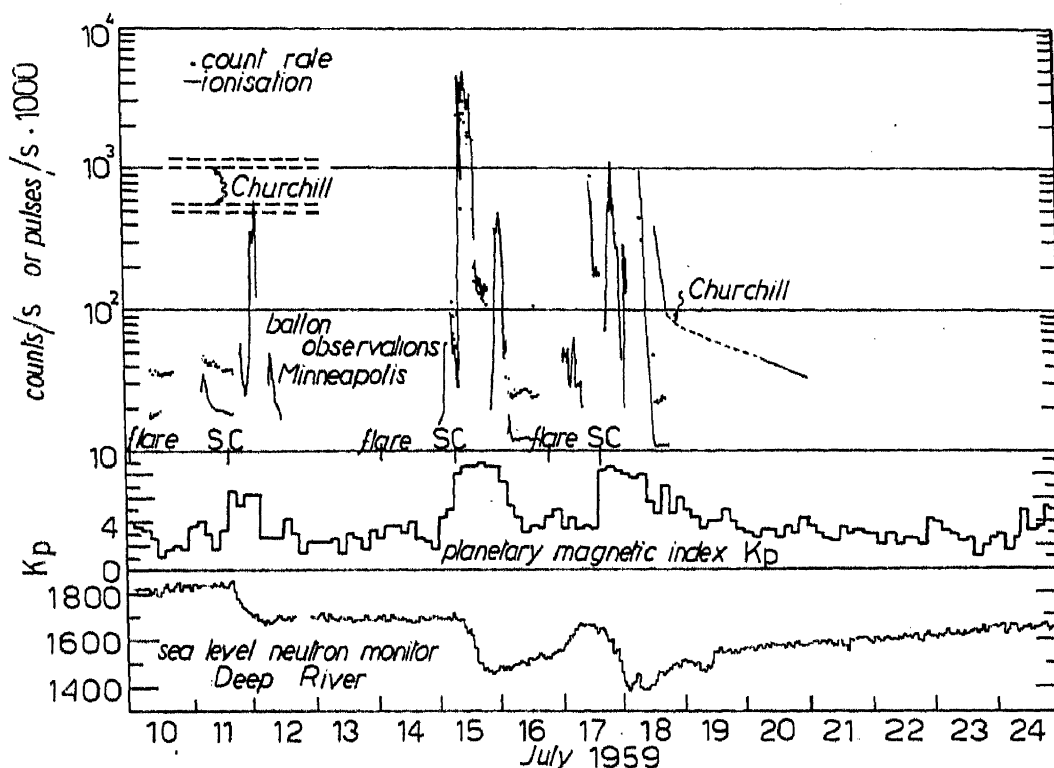


Fig. 15. — Upper: High-altitude total ionization (solid curve) and counter rate (dotted curve) during the three July events. Note logarithmic scale. Center: Note general correlation between influx of radiation and planetary K indices. Lower: Sea-level neutron intensity showing the three large Forbush decreases strongly correlated with the magnetic activity. On July 17, the flare cosmic rays produced a sea-level effect before the geomagnetic disturbance that is attributed to the energy spectrum extending appreciably into the high-energy range of at least 1 GeV. From ref. [16].

storm. This is entirely consistent with the idea that the cut-offs are lowered during the main phase of a geomagnetic storm. When this occurs, auroral X-rays may be observed at balloon altitude. If the balloon is within a few hundred kilometers of a riometer, simultaneous characteristic auroral absorption can often be identified on the riometer record. The energy of the X-rays is in the range of 50 to 100 keV. The electrons required to produce the X-rays are either contained in the plasma cloud which is thought to distort the field

by compression, or they are locally accelerated, or they are dumped from the Van Allen belt. All three possibilities are reasonable, and any or all may occur.

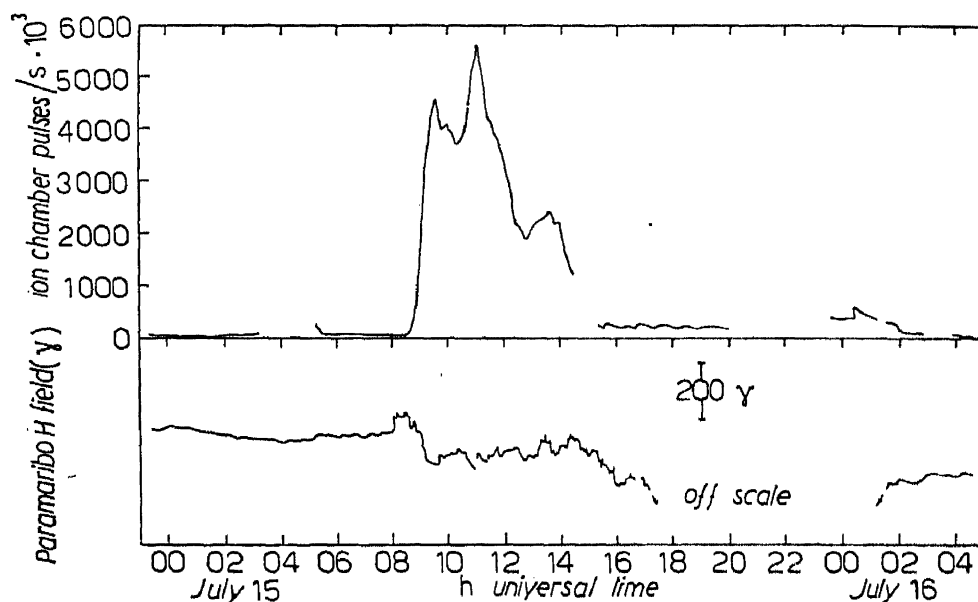


Fig. 16. — High-altitude ionization rate, above, and equatorial horizontal field of the earth below, during the second geomagnetic storm of the July series. Note that the major increase at Minneapolis occurred when the storm field entered the main phase, in this case only about $\frac{1}{2}$ hour after the sudden commencement. From ref. [16].

A theoretical treatment of the generation of X-rays by electrons coming into the earth's atmosphere has recently been given by CLADIS and DRESSLER [18]. The subject of auroral X-rays is an interesting one but it is not considered here [19].

If recovery can be arranged, nuclear emulsions are a good help in the experimental study of solar protons. Although they give no indication of time variations, they allow the identification of small numbers of α 's and heavy nuclei in the flux in a way which is very hard to achieve by any other method. This approach has been exploited by FREIER [20] and FITCHEL [21]. FREIER has used emulsions to study the composition of the incident particles in the manner indicated above, and also to measure the cut-off at Minneapolis as a function of time during the magnetic storm, using emulsion exposed during several flights into the same event.

On March 23, 1958, at 0950 U.T., a class 3+ solar flare was observed with a duration of 248 minutes. A type IV radio noise storm was observed 169 to 9400 MHz and this storm was extremely large at low frequencies. There was a sudden commencement and a Forbush decrease starting late on the 25th, and about this time a P.C.A. event with 10 db of absorption started at Fort Yukon, Canada. Balloon flights were made into this event from Minneapolis on March 21, March 26, and April 8, each carrying a single counter, an ion chamber, and nuclear emulsions. The ion chamber gave a very clear record

of the Forbush decrease at high altitude, agreeing well with the Deep River, Canada, neutron monitor. The increase of ionization which indicates the arrival of the low energy protons was observed directly. The emulsions showed that a large increase in the proton intensity was not accompanied by a corresponding increase in the intensity of α -particles at rigidities greater than 1.3 GV (approximately 200 MeV/nucleon). This puts the proportion α/p above a given rigidity at less than the α/p ratio which is characteristic of galactic cosmic radiation (1/18 in this event). FREIER, BISWAS and STEIN [20] have measured the α/p ratio in the rigidity region 0.9 to 1.3 GV, on September 3, 1960, and find that it is $1/(31 \pm 8)$, characteristic of the sun rather than galactic cosmic rays. The final answer to the question of the composition of the beam is not known, but it can be said that the proportion of α -particles is appreciable and probably variable; their spectrum must, however, wait until the analysis of the November 12, 1960, event is complete.

FICHTEL has studied the proportion of heavy nuclei in the September 3, 1960, solar beam, using the NASA rocket 1019, which was flown at 1408 U.T. (13 hours after the flare). By comparison with nuclear emulsions flown during a test shot at a « quiet » time, he finds that the medium-to-proton ratio above a given rigidity is consistent with the solar constitution as estimated by GOLDBERG [22] and differing from the galactic cosmic ray value by several standard deviations. Statistics were too poor to permit deduction of a spectrum (Tables VII and VIII).

TABLE VII. — *Ratio of nuclei with $6 \leq Z \leq 8$ to those with $Z=1$ in solar beams.*

Charge	Ratio
September 3, 1960, SBE	$(0.8 \pm 0.3) \cdot 10^{-3}$
November 12, 1960, SBE	$(1.9 \pm 0.6) \cdot 10^{-3}$
Sun	$(1.5 \pm 0.5) \cdot 10^{-3}$
Universe	$(1.6 \pm 0.8) \cdot 10^{-3}$
Cosmic rays	$(10 \pm 2) \cdot 10^{-3}$

TABLE VIII. — *Relative abundances of nuclei in solar beams.*

Charge	3 to 5	6	7	8	10	$Z > 10$
September 3, 1960, SBE	0	10	$4 \leq$	11	5	2
November 12, 1960, SBE	0	13	$6 \leq$	23	3	1
Sun	10^{-5}	10	2	18	1	2
Cosmic rays	6	10	$5 \leq$	6	2	5

The Minnesota group has used emulsions to supply additional evidence of the cut-off changes. They estimate in this way that the normal cut-off at

Minneapolis is about 280 MeV, and show how this is altered by a magnetic storm (Fig. 15 and 16).

Apart from the authors already mentioned, R. R. BROWN has made many flights from College, Alaska, which give results in general agreement with those discussed here. Also, the Russians [23] have made a number of balloon flights at various latitudes.

5. - Rocket observations during September and November, 1960.

We shall now turn to the rocket observations of solar protons carried out at Churchill by DAVIS *et al.* [24-26]. These rockets were Nike-Cajun two-stage types and reach a peak altitude of 130 km, spending about 200 s above 90 km. In an attempt to cover as wide an energy and intensity range as possible, a rather complicated system was used. Two scintillation counters and a Geiger counter were carried by the rocket, each telemetering information through a ratemeter. The ratemeter output was arranged to show pulses on the record when the rate was low, and give a voltage output for high rates. In this way, counting rates ranging from a few per second up to 20 000 per second were successfully recorded. Four events were encountered during a continuous watch lasting from June 6, 1960, to the end of the November 12, 1960 event. Arrangements were made with the personnel of the McMath-Hulbert Observatory, the Sacramento Peak Observatory, the Lockheed Observatory, and the University of Hawaii for them to telephone Churchill when a suitable

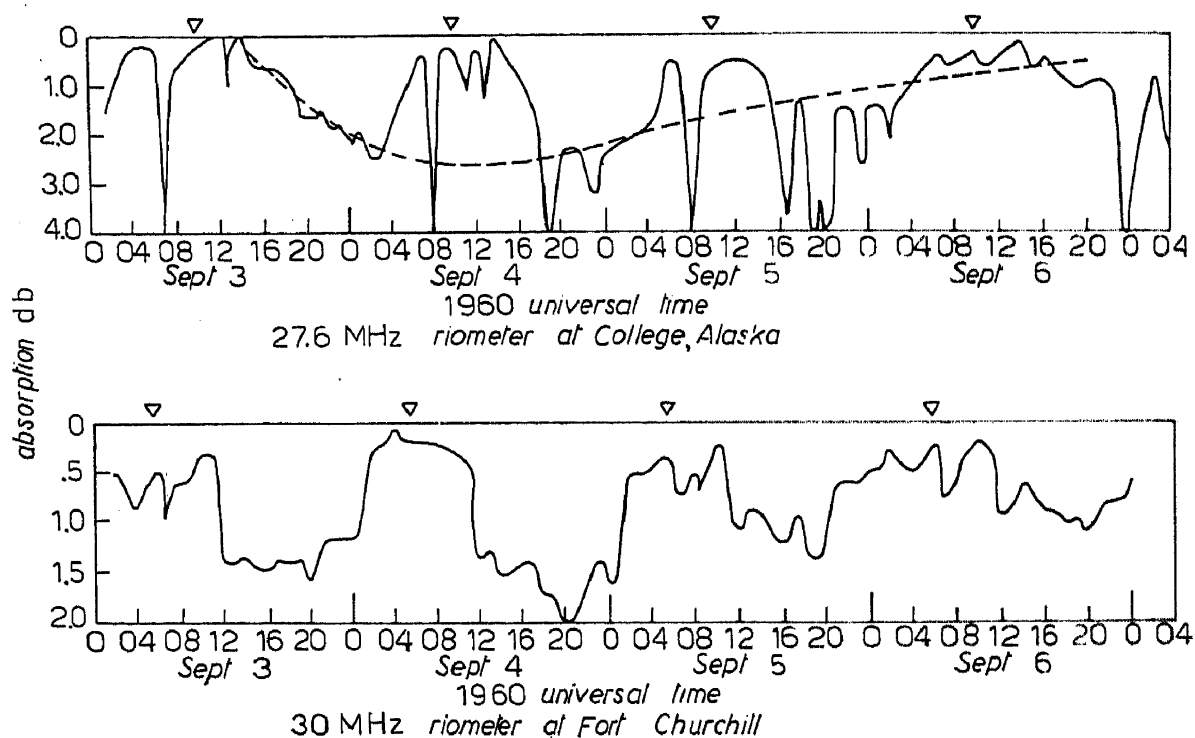


Fig. 17. - Polar cap absorption event of September 3, 1960.

flare was seen. The riometer stations at College, Kiruna, and Churchill were also available by telephone. The rocket instruments were kept in readiness and the rocket was on the launcher at all times. The first event which was detected was on September 3, 1960. The delay between the flare and the start of absorption occurring at 0040 U.T. was about 10 hours, very much longer than usual. The riometer record shown in Fig. 17 and Fig. 1 gives these results to a larger scale, showing the times of the rocket firings, and the diurnal effect. The diagnosing of the rather small absorption during the early part of September 3 was helped by the presence in Churchill of a team from Dr. Winckler's group who were launching balloons. They had a balloon up, and their telescope rate was sufficiently high (25 times normal) to decide the issue. In Table IX, taken from [17], the solar-terrestrial events during the period of interest are shown.

TABLE IX (*). — *Solar, terrestrial time table.*

Time (U.T.) date	Event	Comment on tentative identification
0700, Sept. 2	Class 3 flare	Solar coordinates 19° N, 25° W; Source of magnetic cloud, SC and Forbush decrease
2230, Sept. 2	Class 3 flare	Solar coordinates 21° N, 31° W; Source of apparent non-magnetic cloud but strong magnetic storm
0040, Sept. 3	Class 3 flare	Solar coordinates 17° N, 90° E; Source of cosmic rays
0116, Sept. 3	Class 3 flare	Cosmic ray flare maximum, X-ray burst (local)
0230, Sept. 4	SC	Begin magnetic storm from 0700 flare Sept. 2
0230, Sept. 4	Forbush decrease	Indicates magnetic character of solar cloud
1830, Sept. 4	SC	Begin magnetic storm from 2230 flare Sept. 2 No Forbush effect

(*) From ref. [27].

It is assumed that the third flare is the one of interest. During the early part of the previous night at Churchill, there was an extensive and impressive auroral display, presumably from an earlier flare. Figure 18 shows that the radiation did not reach its maximum intensity until about nine hours after

the flare. The dotted curve 7a represents the spectrum deduced from the emulsions on flight 1019, and curve 8 represents the results from the Geiger counters on flight 1020. The Geiger counter results from flight 1019 agree well with the emulsions but are not shown here to avoid confusion. We note that the spectrum of protons curved away from a power law at low energies,

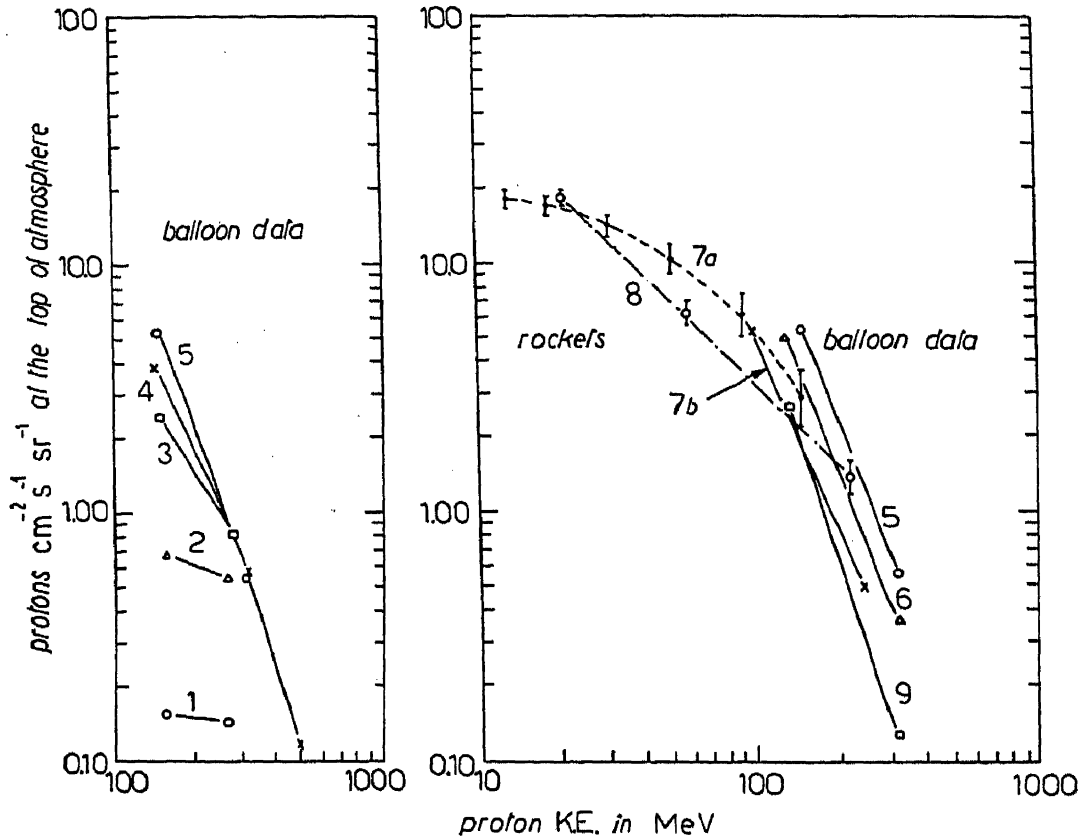


Fig. 18. — Balloon and Rocket observations during September 3, 1960 solar proton event, Churchill. 1) 0400 U.T.; 2) 0600 U.T.; 3) 0800 U.T.; 4) 0900 U.T.; 5) 1000 U.T.; 6) 1200 U.T.; 7a) 1410 U.T. Rocket 1019; 7b) 1410 U.T.; 8) 1730 U.T. Rocket 1020; 9) 2400 U.T.

but that in the region of 100 MeV agreement between balloon and rocket observations is good. Also, the intensity was the same at the time of both firings, and it was probably essentially constant for most of the time between. The spectrum steepens with time during both the increase and the decrease of the intensity. The energy spectrum observed is a result of dispersion operating on the original spectrum, which must be assumed to be injected by the sun during a period of time of the order of one hour, the duration of the radio event. WINCKLER [27] has suggested diffusion with an energy-dependent mean free path, to account for this, consistent with the prompt arrival of particles on February 23, 1956, for example. In passing, it is interesting to note that very low energy particles were seen over Minneapolis, the instrumental limit being

air absorption, on September 4 after the main phase of the geomagnetic storm began (*).

From the September 3 event, the conclusion is that particles with energies down to 13 MeV were detected, that they were protons, and that their distribution in angle above the atmosphere was closely isotropic. It should be pointed out here that D'ARCY [28] has drawn attention to a possible error of interpretation of balloon flight results. The X-rays produced by nuclear reactions in the atmosphere may lead to a larger counting rate than is characteristic of the primary protons. This error could occur in the Geiger counter results given here due to reactions in the (Al) walls of the rocket. However, the agreement between the emulsion and Geiger counter points shows that this effect is small.

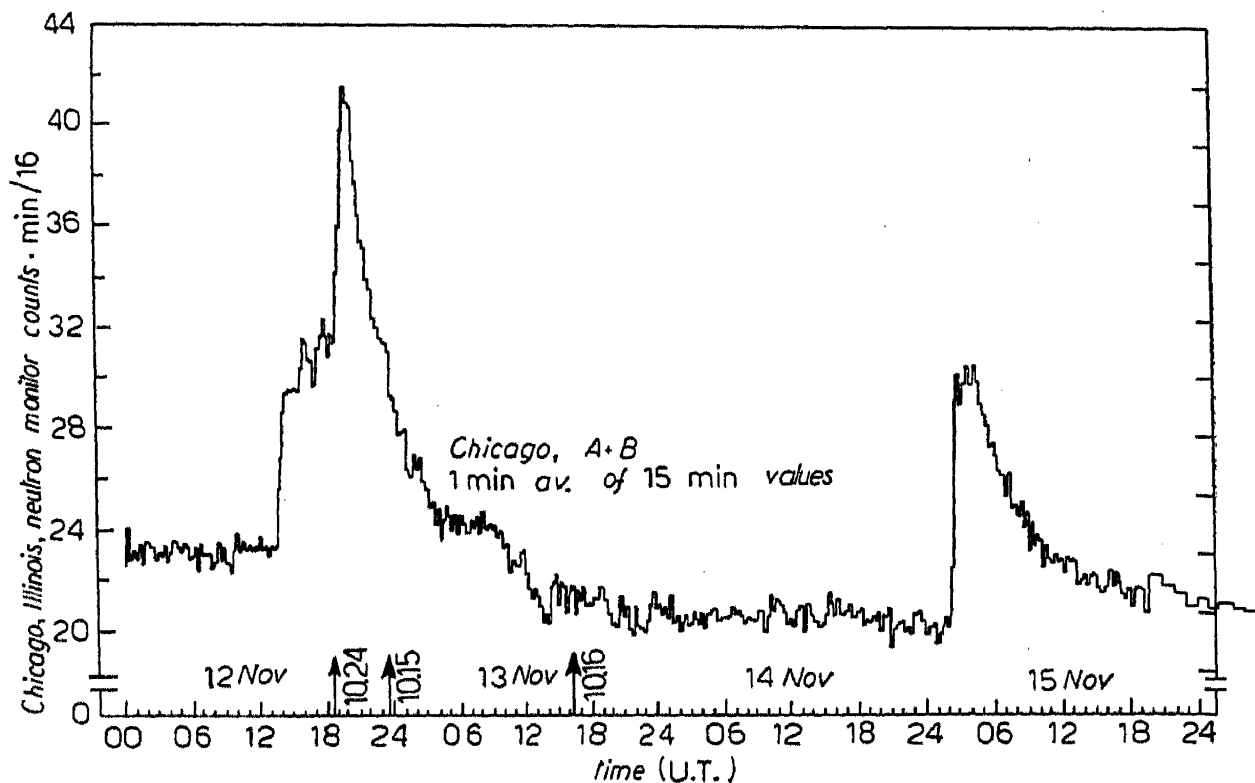


Fig. 19. -- Neutron monitor record from Chicago for November 12-16, 1960.

The solar proton event of November, 1960, was one of the largest yet observed—perhaps it will be the largest in this solar cycle—and NASA was fortunate enough to fire seven rockets at various times during the period of high proton intensity. Analysis of the results from the first three of these, two fired on November 12 and one on November 13, is fairly complete, and the results will be discussed here. For this event we have scintillation counter

(*) Because of a malfunction, the results obtained from the scintillators on September 3 are not presented here.

records that are self-consistent and show the presence of particles with energies down to 2 MeV. After further analysis we shall have information at much lower energies. Figure 19 shows the times of these first three shots, superimposed upon the neutron monitor record from the Chicago station (SIMPSON, private communication). The timing of the first two shots was, to say the least, lucky.

It is necessary now to discuss the method of analysis of the Scintillator results, so that their significance can be appreciated. The two scintillation counters carried by the rockets were fitted with a 0.24 g/cm^2 CsI crystal in one case, and ZnS powder phosphor in the other. The ZnS results have not yet been reduced but they will eventually give information down to 200 keV. The CsI scintillation counter incorporates a rotating switch which continually changes the load of the photomultiplier. The energy loss *vs.* energy curve for the crystal is shown in Fig. 20. Each load resistor corresponds to a different

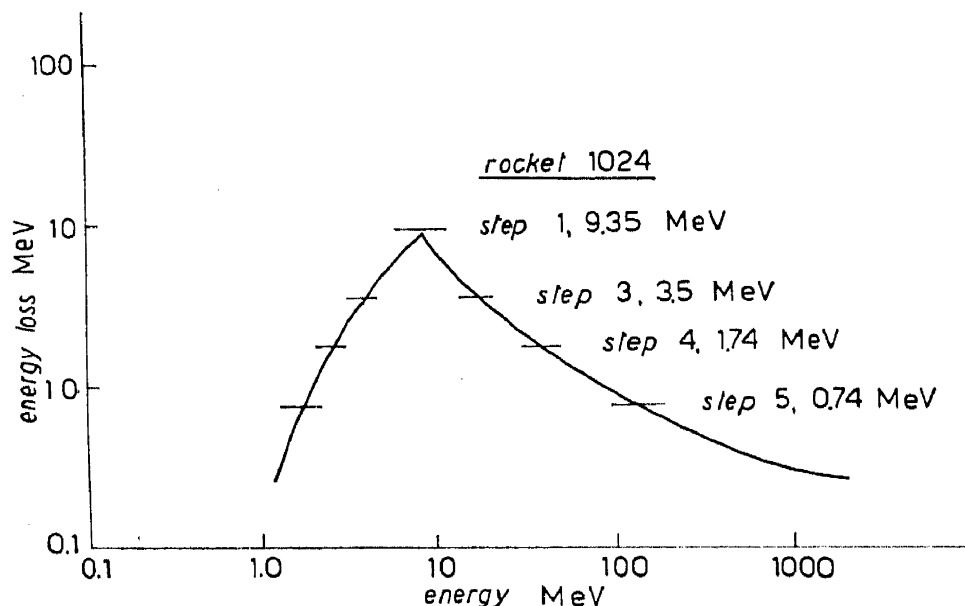


Fig. 20. - Energy sensitivity intervals for CsI scintillation counter.

horizontal line on this diagram the equivalent energy loss being exceeded by protons in a certain energy range. In particular step 1, the uppermost step, would not count at all if it were not for the fact that some protons enter the crystal at oblique angles, outside the proscribed solid angle, and lose more than the 8 MeV characteristic of perpendicular penetration. We have used in this analysis, the ratios between the step rates $\frac{4}{3}$, $\frac{3}{2}$, $\frac{1}{2}$. We start by calculating these ratios assuming protons incident isotropically upon the detector and rocket, and having a power law spectrum $N(>E) = A/E^n$. This calculation was made for a series of values of n , and the results for the first rocket, NASA 1024 rocket, are shown in Fig. 21. The observed ratios are seen superimposed, indicating a best fit for a value of $n = 1.6 \pm 0.1$. The intensity of

protons having energies above 2 MeV can be obtained from the rate on the « widest » step, (step 5) once a slope for the best power law approximation has been determined.

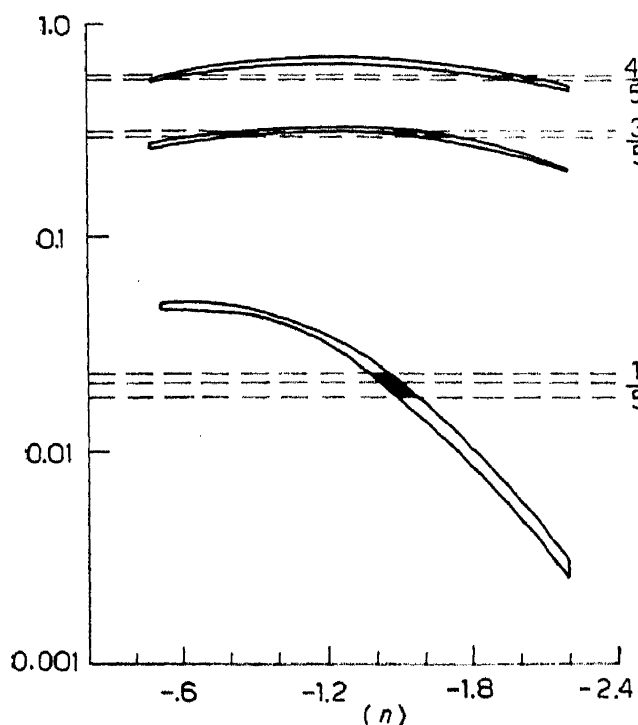


Fig. 21. - Ratio calculation results for NASA rocket 1024, $n = 1.6 \pm 0.1$.

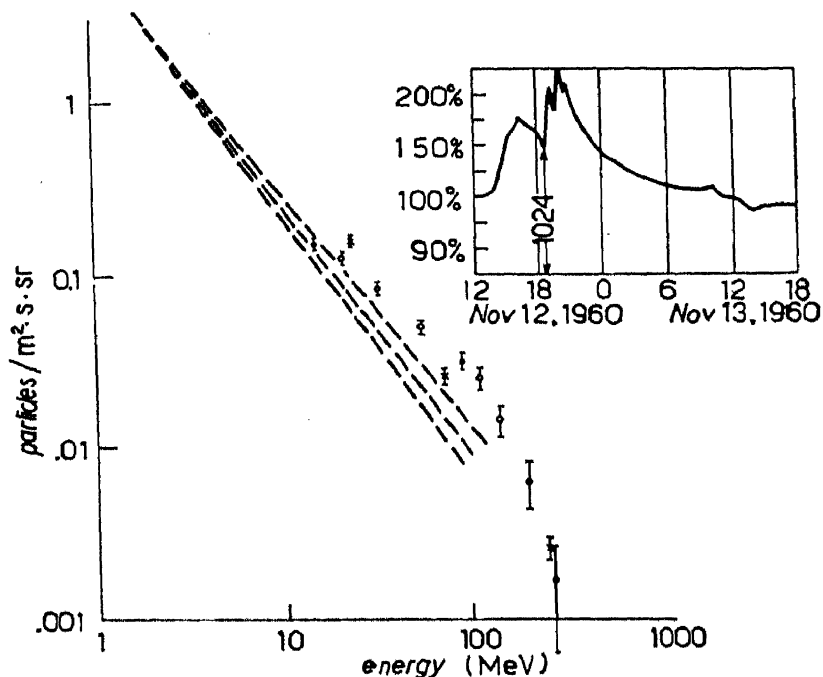


Fig. 22. - Spectrum of protons obtained by NASA rocket 1024, $n = 1.7 \pm 0.2$. $\circ$ emulsion points.

Thus, in each case the results from the CsI scintillation counter determine the intensity above 2 MeV and the slope of the best straight line approximation to the spectrum between there and about 100 MeV (Fig. 22). We do not believe firmly in the existence of a simple power law spectrum in this region, but this approximation can be used until more definite evidence is available. In particular, the emulsion results presented by FITCHEL [29] indicate appreciable flattening of the spectrum in the region between 100 and 15 MeV. This should not be thought to be contradictory to the results presented here.

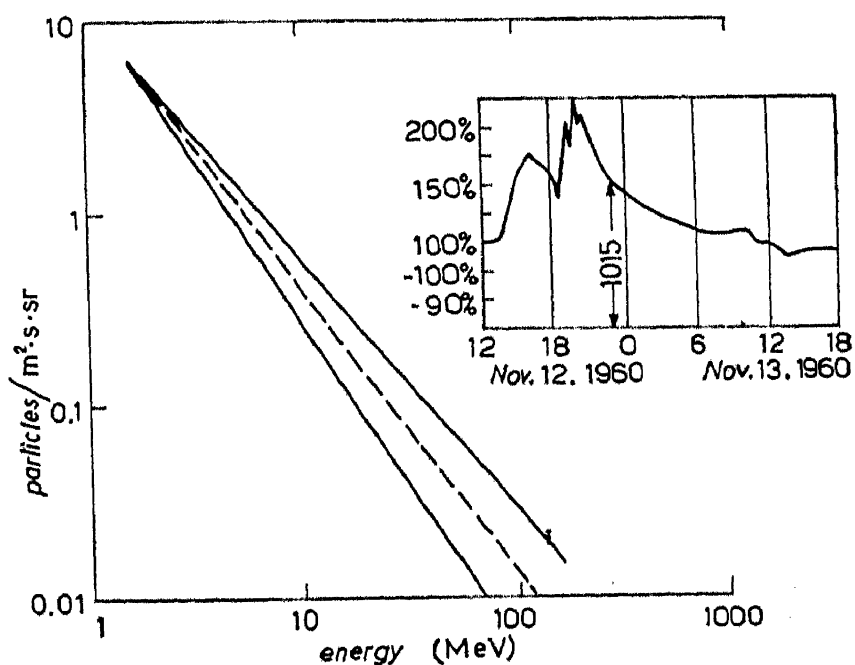


Fig. 23. - Spectrum of protons obtained by NASA rocket 1015, $n = 1.5 \pm 0.2$.

If a spectrum deduced by some other means becomes available, it will not be difficult to test the degree to which it is compatible with these results.

As examples of this procedure Fig. 22, 23 and 24 show the actual results of rockets 1024, 1015, and 1016.

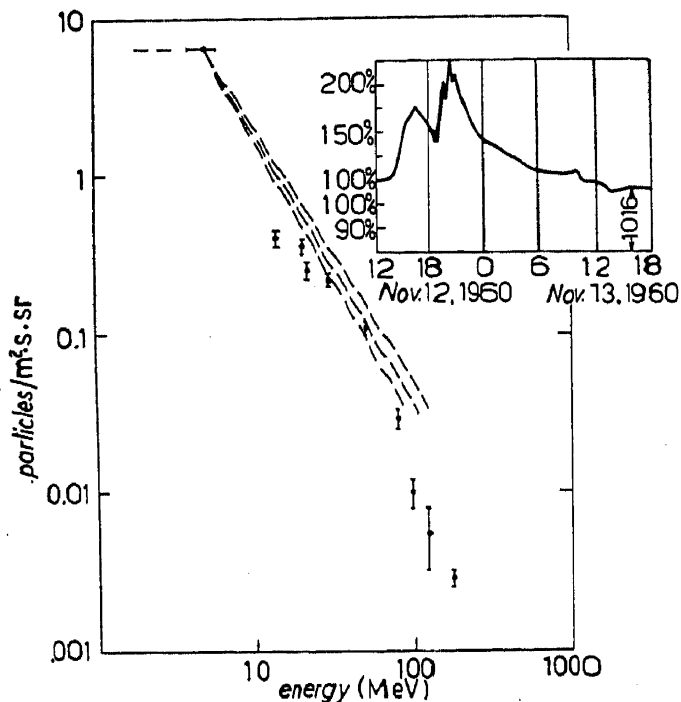


Fig. 24. - Spectrum of protons obtained by NASA rocket 1016. • emulsion results.

The inset shows the timing of these shots compared with the neutron monitor increase observed at Deep River. Evidence that the protons arrive isotropically in the upper hemisphere at the top of the atmosphere is provided by the angular distributions shown in Fig. 25 and 26.

If the ratio $\frac{4}{3}$ is closely equal to the ratio $\frac{3}{2}$, this indicates both a steep spectrum, and an absence of low energy particles. This situation occurs with Rocket 1016 (Fig. 27), and this is dealt with by cutting off the spectrum at a low energy and making a two-parameter fit. It was found that the best results are obtained by assuming no particles with energies below (5 ± 1) MeV.

Thus, to the degree of accuracy of our experiment, the proton spectrum can be represented by a power law, with or without a cut-off above 2 MeV, and an intensity value either at 2 MeV or at the cut-off energy, deduced from the rate of step 5. Since the analysis of this event is as yet very incomplete, no interpretation will be presented here. However, some remarks can be made.

First of all, the neutron monitor increase represents the way in which the intensity of particles at energies about 1 GeV behaved in time. There is a fairly rapid increase, a «step», and a further increase followed by a slower decay. This «step» feature will be discussed later. The intensity of protons at low energies (3 MeV) had the values shown in Table X. This table

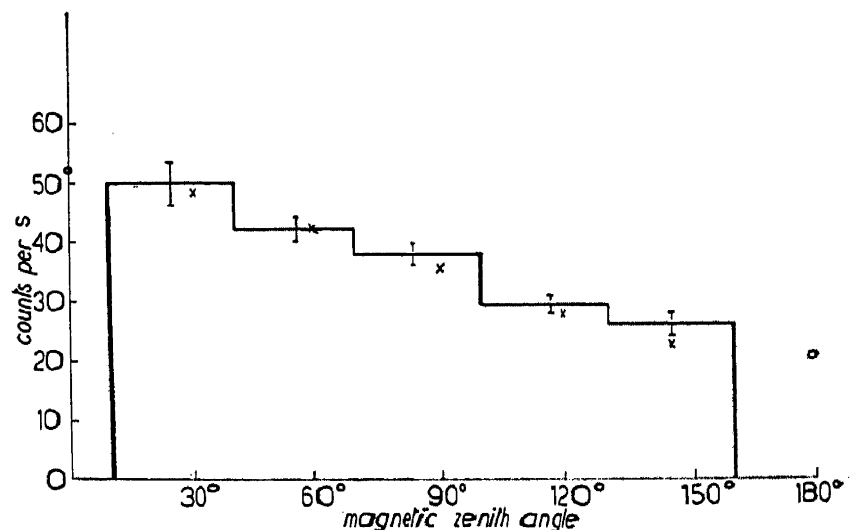


Fig. 25. - Geiger counter Zenith angle dependence. x calculated assuming isotropy. NASA rocket 1020.

illustrates the steepening of the spectrum, since the neutron monitor increase had entirely disappeared by the time of rocket 1016.

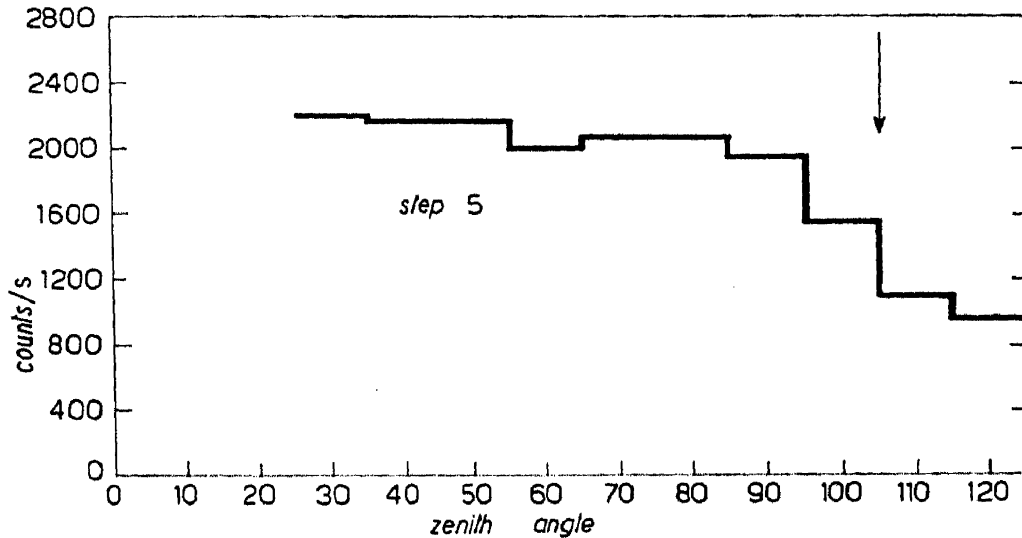


Fig. 26. - Angular distribution of protons with energies above 2 MeV, from the CsI scintillation counter in NASA rocket 1024.

TABLE X. - *Intensity of protons with energies greater than 2 MeV during the November, 1960, event.*

Time (U.T.), date	Rocket number	Intensity above 2 MeV $\left(\frac{\text{particles}}{\text{cm}^2 \text{ s sr}}\right)$
1648, Nov. 12	1024	30 500
2330, Nov. 12	1015	62 000
1600, Nov. 13	1016	63 500

The absence of particles with energies below (5 ± 1) MeV found by rocket 1016 is a very interesting result, and may be connected with the fact that the rocket was fired immediately after the large negative excursion of the magnetic field observed at Fredericksburg, Virginia (see Fig. 28).

The November event can be interpreted within the framework of the ideas of GOLD [29]. In his view a disturbance on

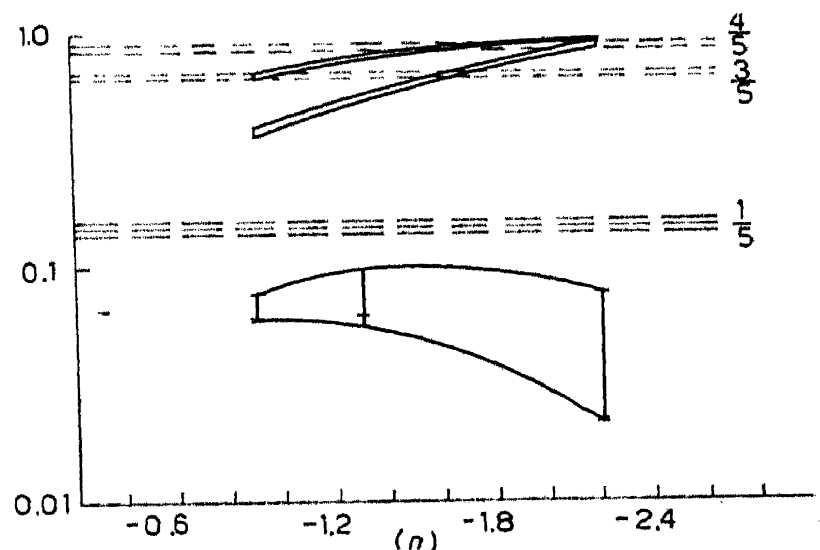


Fig. 27. - Ratio calculations results for NASA rocket 1016. 5 MeV cut-off, $n = 1.7 \pm 0.1$.

the sun, not necessarily one resulting in the acceleration of particles, may drag out the lines of force of the sun's field, forming a tongue, or «bottle». If the acceleration process occurs at the point of origin of this elongated field, particles can be trapped within it for some time, and if the earth enters the

region, a solar proton event is observed. If acceleration occurs at some other point, then any particles which can enter the region may be trapped by it, but first they must travel in the local field region of the sun to the base of the bottle.

On this view there was a bottle, caused by an earlier flare, on the way toward the earth on

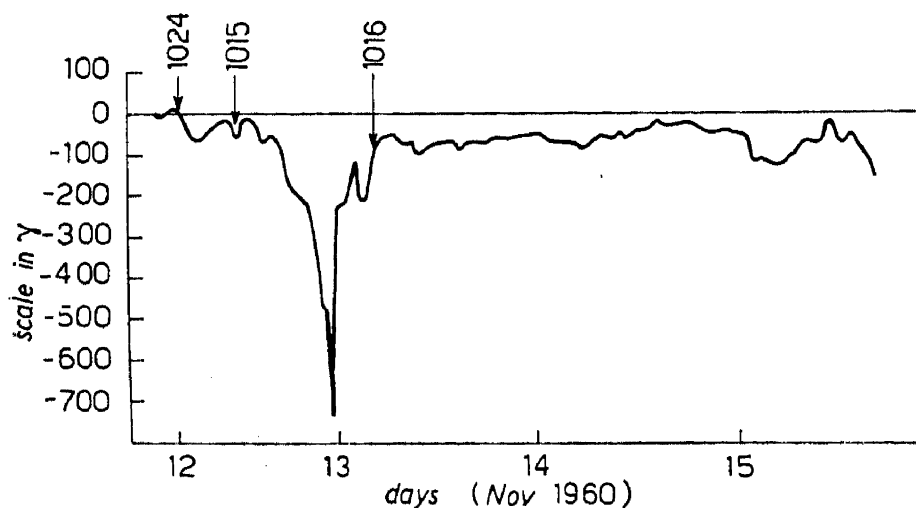


Fig. 28.—Magnetic observations during November, 1960, at Fredericksburg, Virginia.

November 2. When particle acceleration took place, the initial increase in particle intensity observed represents leakage from the bottle, since the earth was not then inside it. The «step» represents the entry of the earth; and a corresponding increase in intensity took place. The Forbush decrease occurring at the time of the step, confirms the correctness of this general picture, since it is supposed to occur when the earth is shielded from galactic radiation by the magnetic walls of the bottle. As McCracken [30] and others have shown, these ideas are capable of introducing a considerable simplification into the classification of cosmic ray increases. «Sharp risers», like the May 4, 1960, event where the intensity increases in a very short time and where there are marked impact zones and anisotropy, represent cases where the earth is already inside the bottle at the time of the increase. «Slow risers», which show isotropic intensity and absence of impact zone effects, correspond to situations in which the protons must diffuse out of the walls of the bottle before reaching the earth.

It can be demonstrated that the impact zones which occur in the first situation correspond to a source direction about 50 to 60 degrees west of that of the sun, which is thought to be due the dragging of the field by solar rotation. The principal objection which can be raised to these ideas is that of PARKER, who finds it difficult to supply the energy necessary to project the field, which must travel in the form of a shock wave.

REFERENCES

- [1] R. R. BROWN and R. A. WEIR: *Ark. Geofys.*, **3**, 523 (1961). See other useful papers in this issue.
- [2] G. C. REID: *Journ. Geophys. Res.*, **66**, 4071 (1961).
- [3] C. G. LITTLE and H. LEINBACH: *Proc. I.R.E.*, **46**, 334 (1958).
- [4] L. HERZBERG: *Journ. Geophys. Res.*, **65**, 3505 (1960).
- [5] D. K. BAILEY: *Proc. I.R.E.*, **47**, 255 (1959).
- [6] C. COLLINS, D. H. JELLY and P. T. MATTEWS: *Canad. Journ. Phys.*, **39**, 35 (1961).
- [7] M. R. KUNDU and F. T. HADDOCK: *Nature*, **186**, 610 (1960).
- [8] J. F. DENISSE, A. BOISCHOT and M. PICK-GUTMANN: *Propriétés des éruptions chromosphériques associées à la production de rayons cosmiques par le soleil*, in: *Space Research, Proceedings of the First International Space Science Symposium*, Nice, France, January 1960 (Amsterdam, 1960), p. 637.
- [9] J. A. SIMPSON: *Suppl. Astrophys. Journ.*, **4**, 378 (1960).
- [10] T. OBAYASHI and Y. HAKURA: *Enhanced ionization in the polar ionosphere and solar corpuscular radiation*, in: *Space Research, Proceedings of the First International Space Science Symposium*, Nice, France, January 1960 (Amsterdam, 1960), p. 665.
- [11] H. W. DODSON and E. R. HEDEMAN: *Journ. Geophys. Res.*, **65**, 123 (1960).
- [11b] C. S. WARWICK and M. W. HAURWITZ: *Journ. Geophys. Res.*, **67**, 1317 (1962).
- [12] K. A. ANDERSON: *Preliminary Study of Prediction Aspects of Solar Cosmic Ray Events*, NASA Technical Note TN D-700 (April, 1961).
- [13] F. BACHELET, P. BALATA, A. M. CONFORTO and G. MARINI: *Nuovo Cimento*, **16**, 292 (1960).
- [14] T. OBAYASHI: *Rpt. of Ionosphere and Space Res. in Japan*, **13**, 201 (1959).
- [14b] T. OBAYASHI: *Journ. Geomag. and Geoelec.*, **13**, 26, (1961).
- [15] P. ROTHWELL: *Journ. Geophys. Res.*, **64**, 2026 (1959).
- [16] J. R. WINCKLER: *Journ. Geophys. Res.*, **65**, 1331 (1960); J. R. WINCKLER, P. D. BHAVSAR and L. PETERSON: *Journ. Geophys. Res.*, **66**, 995 (1961).
- [17] K. A. ANDERSON, R. ARNOLDY, R. HOFFMAN, L. PETERSON and J. R. WINKLER: *Journ. Geophys. Res.*, **64**, 1133 (1959).
- [18] J. B. CLADIS and J. A. DRESSLER: *Journ. Geophys. Res.*, **66**, 343 (1961).
- [19] K. A. ANDERSON: *Journ. Geophys. Res.*, **65**, 551 (1960); K. A. ANDERSON and D. C. ENEMARK: *Journ. Geophys. Res.*, **65**, 3521 (1960).
- [20] S. BISWAS P. S. FREIER, E. P. NEY and W. STEIN: *Midwest Cosmic Ray Conference Papers* (unpublished).
- [21] C. E. FICHEL and D. E. GUSS: *Phys. Rev. Lett.*, **6**, 495 (1961).
- [22] L. GOLDBERG, E. A. MULLER and L. H. ALLER: *Suppl. Astrophys. Journ.*, **5**, 1 (1960).
- [23] A. N. CHARAKHCHIAN, V. F. TULINOV and T. N. CHARAKHCHIAN: *Cosmic rays emitted by the sun*, in: *Space Research, Proceedings of the First International Space Science Symposium*, Nice, January 1960 (Amsterdam, 1960), p. 649.
- [24] L. R. DAVIS, C. E. FICHEL, D. E. GUSS and K. W. OGILVIE: *Phys. Rev. Lett.*, **6**, 492 (1961).
- [25] K. W. OGILVIE, D. A. BRYANT and L. R. DAVIS: *Journ. Geophys. Res.*, (March, 1962).

- [26] L. R. DAVIS and K. W. OGILVIE: *Journ. Geophys. Res.* (to be published).
- [27] J. R. WINCKLER: *Phys. Rev. Lett.*, 6, 488 (1961).
- [28] R. G. D'ARCY: *Balloon Observations of Solar Protons in the Auroral Zone*, July 1959, Doctorial Thesis (Berkeley, December 1960).
- [29] T. GOLD: *Suppl. Astrophys. Journ.*, 4, 406 (1960).
- [30] K. G. McCracken: *Journ. Geophys. Res.*, 67, 423, 435, 447 (1962).

Propagation of Solar Particles and the Interstellar Magnetic Field.

C. S. WARWICK

National Bureau of Standards, Central Radio Propagation Laboratory - Boulder, Colo.

A study of solar activity associated with polar cap absorption (PCA) events has revealed some characteristics that bear on the propagation of these particles in an interplanetary field. These characteristics are tabulated below, along with corresponding properties of solar particles in higher and lower energy ranges.

Ground-level increases 10^9 to 10^{10} eV	PCA 10^7 to 10^8 eV	Geomagnetic storm 10^4 eV
Not associated with flares on the far-side of the sun	No far-side flares	No far-side flares
Significant western excess of associated flares	Slight western excess	No asymmetry
Stopped at maximum solar activity	Delayed, but not stopped at maximum	No effect

Solar particle emission in each energy range is strongly associated with activity on the visible hemisphere of the sun, showing that in general no efficient mechanism operates to bring particles to the earth from the far side of the sun, that is, there is no reflecting magnetic shell and no strongly curved spiralled lines of force in the solar equatorial plane. There is a significant excess of western flares associated with high-energy solar particles; this effect is lacking, or at least much less pronounced for the particles of lower energy. The frequency of occurrence of PCA events relative to solar activity such as large flares and radio bursts remains constant throughout the solar cycle,

while the relative frequency of occurrence of ground-level events is low near the maximum solar activity. In the latter type of event, the particles may be emitted from the sun at solar maximum, but prevented from reaching the earth. The medium energy particles that cause PCA apparently are nearly always able to reach the earth, but the delay time between the solar flare and onset of absorption at the earth may be several hours near solar maximum while away from maximum, delay times of several minutes, appropriate to the particle energy, are sometimes observed. The lower energy particles associated with geomagnetic disturbance do not show any effect of this kind. The information in the table may be summarized by saying that solar particles show many evidences of the effects of magnetic fields between the earth and the sun, and that the higher the energy of the individual particles, the larger is the effect.

If these particles reached the earth as individual particles, whether along relatively simple trajectories in the magnetic field, or by a diffusion-like process, or by guidance along a line of force connecting earth and sun, the magnitude of the gyro-radius of the particle would indicate how easily the particle moves in the field. We should expect that the higher the energy of the particle, the more frequently and promptly these particles would reach the earth. The fact that we observe rather that the low-energy particles reach the earth most easily suggests that in the two lower energy ranges the particles are acting, not as single particles, but as a group. Then the relevant measure is the energy density of the group of particles. Because the energy spectrum of these solar emissions is steep, that is, the number of particles having a given energy increases rapidly as individual particle energy decreases, the energy density is highest for the low energy particles. If we accept the idea that in PCA events the particles are able to move through the interplanetary magnetic field by virtue of the fact that their energy density is greater than that of the field, we can estimate the range of values of magnetic field strength that would be effective in delaying these particles. One finds that a field of 10^{-5} to 10^{-4} G would balance the energy density of the solar particles in PCA events.

Such a field would prevent the passage from sun to earth of the high energy solar particles associated with ground-level events. Let us assume that at times near minimum solar activity, the interplanetary magnetic field is identical with a uniform galactic arm field directed towards *Cygnus* and so at an angle of about 60° to the solar equatorial plane. If solar particles of energy 10^9 eV are to reach the earth in one half a revolution, the strength of this field must be no larger than 10^{-6} G. If such a field were directed northward, particles emitted radially from the sun would reach the earth more easily from the west limb of the sun than from the east limb.

When low-energy particles of the sort associated with geomagnetic distur-

bance are emitted by the sun, they will push this field ahead of them and form a region of higher intensity. Such regions might have field strength of 10^{-5} to 10^{-4} G, and be numerous enough in the region from $\frac{1}{2}$ A.U. to 10 A.U. from the sun that particles traveling from sun to earth would be almost certain to encounter at least one such region near solar maximum. These regions would tend to deflect the high energy particles away from the ecliptic plane, so that they would fail to reach the earth. The lower energy particles associated with PCA might, because of broader energy spectrum and higher density, overcome the field after some delay.

Regions of enhanced magnetic field moving outwards from the sun provide a possible explanation of the main features of solar modulation of galactic cosmic rays. When a cosmic-ray particle encounters such a moving cloud, the effect of the encounter is to add to the particle velocity the velocity of the cloud. Although the cloud velocity is only about 1% of the particle velocity, the fact that it is systematically outward makes even this very small velocity effective in lowering the density of cosmic-ray particles in the vicinity of the sun. If the particle velocity distribution is originally isotropic, collisions with the moving clouds tend to make the velocity distribution isotropic in a reference system in which the clouds are at rest, or, in the original reference frame, to introduce an outward streaming.

We may investigate the effect of this added velocity on the cosmic-ray flux observed at the earth. The equation of continuity in cylindrical co-ordinates for the case when the only component of velocity is radial is:

$$\frac{\partial N}{\partial t} = -\frac{1}{r} \frac{\partial}{\partial r} [rNV].$$

We expect that the solar clouds attain distances of about 10 A.U. or less, so that beyond this distance, the systematic velocity V falls to zero. If we simply consider the decrement in the quantity $[rNV]$ between 1 A.U. and 10 A.U., we find

$$-\frac{\partial N}{\partial t} \sim \frac{\Delta[rNV]}{r\Delta r} = \frac{1 \text{ A.U.} \times N \times V}{5.5 \text{ A.U.} \times 9 \text{ A.U.}},$$

or

$$-\frac{1}{N_0} \frac{\partial N}{\partial t} = \frac{NV}{N_0 \times 49.5 \times 1.5 \times 10^{13}} \text{ s}^{-1}.$$

Taking a value for the solar-cycle variation in cosmic-ray flux of

$$-\frac{1}{\Phi} \frac{\partial \Phi}{\partial t} = -\frac{1}{N} \frac{\partial N}{\partial t} = -60\% \text{ in 4 years} = 0.475 \times 10^{-8} \text{ s}^{-1}, \quad \text{and} \quad \frac{N}{N_0} = .40,$$

we find that $V = .9 \cdot 10^7 \text{ cm s}^{-1}$. This seems to be a reasonable value, since the upper limit for the systematic velocity will be the cloud velocity of $\sim 10^8 \text{ cm s}^{-1}$.

The systematic velocity itself gives a diurnal variation in cosmic-ray flux. Coming from the direction of the sun, the observed flux is $N(V_p + V)$, where V_p is the original velocity of the particle, and V is the systematic velocity. On the night side one sees $N(V_p - V)$; so that the amplitude of the variations is

$$\frac{\Delta\Phi}{\Phi} = [N(V_p + V) - N(V_p - V)]/N V_p = \frac{2V}{V_p}.$$

If $V_p = 2 \cdot 10^{10} \text{ cm/s}$ and $V = 10^7 \text{ cm/s}$, $\Delta\Phi/\Phi = 10^{-3}$.

The region of enhanced field may act as a barrier for cosmic-ray particles, causing locally a Forbush decrease in a way similar to that suggested by MORRISON. Comparison of variations of the geomagnetic field with measures of the interplanetary magnetic field made from Pioneer V suggests that the thickness of the cloud front is about 10^{11} cm . This thickness is consistent with the relative inefficiency of the shielding, the rapidity of the Forbush decrease, and the association of the decrease with geomagnetic disturbance with sudden commencement. For particles with high enough energy that their gyro-radius is large compared to the dimensions of the cloud, the effect of the encounter on the particle motion will be small. The high-energy particles will be affected less than the low-energy particles, so that all of the solar modulation effects would be energy-dependent. Thus, the systematic velocity introduced into the cosmic-ray velocity distribution by encounters with outward-moving solar clouds provides directly an explanation of the diurnal variation of cosmic-ray flux, the shielding effect of the cloud may explain the Forbush decrease, while the systematic outward velocity results in a decrease in density that accounts for the 11-year modulation.

Solar Corpuscular Radiation and the Nature of the Earth's Immediate Interplanetary Environment.

E. R. HARRISON

Atomic Energy Research Establishment - Harwell, Berkshire

1. - Introduction.

Our knowledge of the Earth's immediate interplanetary environment is in an early stage of development. As yet we do not understand the essential physical processes, correlated with solar activity, that govern the phenomena in the neighbourhood of the Earth within distances of 20 Earth radii or more. The intricate nature of the subject stems from the Earth's possession of a magnetic field and from our knowledge that the Sun is a copious emitter of energetic streams of charged particles. The interactions of the solar corpuscular streams impacting against, flowing round, and penetrating into the geomagnetic field are exceedingly complex. Their elucidation is one of the main challenges to the physics of ionized gases.

In our discussion we shall briefly touch on only a few of the problems concerning the environment of the Earth. In Section 2 some of the evidence is given for believing that the Sun is an occasional emitter of streams of charged particles. This is followed in Section 3 with a review of the evidence for and against the solar wind theory. The evidence as it stands is not adverse to the theory, provided one assigns to it a density of less than 100 protons cm^{-3} .

In Section 4, and the subsequent Sections it is assumed that the solar wind has a density of 50 protons cm^{-3} , and the consequences of this assumption are examined. Neither the solar wind nor the occasionally more intense corpuscular streams have any profound effect upon the inner regions of the geomagnetic field which extend symmetrically about the Earth to a distance of normally about 6 Earth radii. This inner region co-rotates with the Earth. The distant geomagnetic field, consisting of the flux-lines which intersect the Earth's surface at all latitudes above the auroral zones, is distributed in a completely asymmetric manner about the Earth. The distant magnetic field

is parted and dragged round to the night side of the Earth to form a « magnetic arch ». The region about the Earth containing the geomagnetic field, referred to as the magnetosphere, has the shape of an elongated tear-drop with its tail at a distance of 20 or more Earth radii.

In Section 5, the origin of auroral streams is discussed, and emphasis is placed on the discovery by radio studies that the aurorae are a night-side phenomenon. Unlike the inner region of the geomagnetic field, the magnetic arch does not rotate with the Earth. Therefore, between the two regions, on the night side of the Earth, there exists a surface of shear which connects directly with the auroral zone. Within this surface of agitated plasma and magnetic fields it is possible that protons (or other ions) are accelerated by Fermi interactions. Alternatively, since the flux-lines within the magnetic arch move around each in synchronism with the Earth's rotation, their length is continually changing in a periodic manner. Protons entrained within the tubes of force are accelerated and decelerated in a similar cyclic manner, and have their maximum energy of about 100 keV above the auroral zones on the night side of the Earth. An acceleration mechanism for electrons is also proposed, and this follows from the electric fields created by the deformation of the polar magnetic flux into helical-like paths. All three of these acceleration mechanisms follow naturally from an asymmetric distribution about the Earth of the magnetic field which connects with the polar regions, and all have the property that they account for auroral displays occurring on the night side of the Earth.

In Section 6 the nature of the gegenschein is discussed, and the possibility that it is located either in the tail-end of the magnetosphere or within an extended « gas-magnetic tail » is examined.

2. – Solar corpuscular streams.

From the world-wide disturbances of the geomagnetic field we know by inference that the Sun is an occasional emitter of corpuscular streams. Shortly after the discovery of the 11 year sunspot cycle LAMONT and SABINE independently discovered in 1851 the 11 year periodic fluctuation in the magnetic declination. In 1859 CARRINGTON [1] observed that a short-lived bright spot on the face of the Sun was later followed by auroral displays and a geomagnetic storm, and was the first to suggest that specific solar and terrestrial events have a common relation. The recurrence of magnetic disturbances every 27 days, in synchronism with the Sun's period of rotation, was first proved conclusively by CHREE and STAGG [2]. From statistical investigations (GREAVES and NEWTON [3, 4]) we now know that there is a definite tendency of the greater geomagnetic storms to follow 1 day after the central meridional

passage of a large sunspot group, particularly when it is accompanied by intense chromospheric eruptions, or « flares ». Smaller magnetic storms are not significantly associated with sunspots but none the less they show a marked tendency to recur every 27 days.

The transit or delay time T , between the central meridional passage of a specific solar event and the onset of a magnetic storm, varies between approximately 15 and 50 hours. In the range of transit times of 21-41 hours magnetic storms are roughly of equal probability; the maximum frequency occurs for transit times of 17-21 hours (NEWTON and JACKSON [5]). One readily finds that protons travelling from the Sun to the Earth in a time T , must have an energy of

$$(1) \quad E = 1 \cdot 10^7 T^{-2} \text{ eV},$$

where T is measured in hours. A transit time of 20 hours therefore implies that the protons arriving at the Earth have a mean energy of 25 keV. The corresponding speed is

$$(2) \quad V = 4 \cdot 10^4 T^{-1} \text{ km s}^{-1},$$

or 2000 km s⁻¹. Doppler shifted hydrogen lines were first observed by MEINEL [6] with a spectrograph pointed towards the magnetic zenith during a brilliant aurora. The violet wing of the H_α profile indicated incoming proton velocities of at least 3000 km s⁻¹.

There is little doubt that the Earth is bathed with a periodic and an irregularly fluctuating solar corpuscular radiation, of which protons have typical energies on the tens of kilovolts range. This conclusion follows from convincing and well established evidence.

3. – Is there a solar wind?

In 1953 STOREY [7] deduced from the propagation characteristics of whistlers that the electron density at a distance of 2-3 Earth radii was approximately 400 cm⁻³. Following Storey's original estimates a number of other studies of different phenomena have yielded a similar density extending for a very much large distance from the Earth. In the same year BEHR and SIEDENTOPF [8] made the first accurate measurements of the degree of polarization of the zodiacal light. By assuming that dust scattering could not account for the polarization effect they estimated that the electron density in the plane of the ecliptic was 600 cm⁻³ at a distance of 1 A.U. from the Sun.

More recently CHAPMAN [9, 10] has drawn attention to the possibility that the Earth's orbit lies within the extended solar corona. He considers a static,

spherically symmetric corona in which heat is conducted radially outwards, and shows that the temperature obeys the law $T \propto r^{-\frac{2}{3}}$. It follows that at the orbit of the Earth the solar gas is at a temperature of about $2 \cdot 10^5$ °K. From the hydrostatic equation one then finds that for a coronal density of 10^8 cm⁻³, the density at the Earth's orbit is of the order of 10^3 cm⁻³. This idealized static model serves the purpose of showing that the solar atmosphere does not have an abrupt fall off in density, and therefore, whatever state it might be in, one might expect it to extend for large distances throughout the solar system.

In the meantime progress was made in a different direction. BIERMANN [11-14] had advanced the hypothesis that in addition to the irregular and periodic corpuscular emissions there is a continuous and permanent flow of ionized gas away from the Sun. The evidence for this view comes from the radial accelerations observed in certain comet tails (BOBROVNIKOV [15], WURM [16]). The repulsive force away from the Sun is occasionally about 100 times the gravitational attraction, and sometimes is even several thousand times as great. From such data BIERMANN estimates that the ionized gas streaming away from the Sun has typically a density of 100 cm⁻³ and a velocity of 500 km s⁻¹ at the distance of the Earth's orbit. Under conditions of enhanced and violent solar activity it is supposed that the density increases by two and even three orders of magnitude, and the velocity increases to possibly $3\,000$ km s⁻¹. KIEPENHEUER [17] discusses and reviews these results and refers to the ionized gas as a « wind » blowing away from the Sun.

PARKER [18, 19] has shown that there is no static solution for the barometric equation in which the hydrostatic pressure of the solar corona goes to zero at infinity, and the corona must therefore continuously expand hydrodynamically away from the Sun. On the basis of this model Parker deduces that at the Earth's orbit the density and velocity is similar to that estimated by BIERMANN. The radial flow of ionized gas away from the Sun is referred to by PARKER as the « solar wind ».

Thus, the picture that emerges is that under quiescent conditions there is an interplanetary medium of ionized gas, and this medium has a density at the Earth's orbit, variously estimated by different authors, of several hundred ions per cm³, and in all probability it possesses a radial velocity away from the Sun of several hundred kilometers per second. Possibly one of the strongest pieces of evidence in support of the solar wind theory comes from the fact that weak aurorae are observed on almost every clear night somewhere in the auroral zones (HARANG [20]). These aurorae appear to have no conspicuous relation with any specific solar event.

However, it would appear that the solar wind theory is not founded on evidence which is altogether secure, and there are many people who yet remain to be convinced of its reality. BLACKWELL and INGHAM [21] have recently

reviewed the evidence for a solar wind and have emphasized that this evidence is still in an unsettled state. BLACKWELL and INGHAM's [21] observations of the zodiacal light in 1958 at the high altitude station of Chacaltaya and the study of the spectra that were obtained, have failed to detect any electrons in the interplanetary region. They show that the upper limit for the electron density at 1 A.U. is 120 cm^{-3} , and that the polarization of the zodiacal light cannot be attributed to the presence of electrons, as was done by BEHR and SIEDENTOPF. Also, STOREY's estimates of the electron density were uncertain and depended markedly on the density distribution close to the Earth. More precise knowledge of the density distribution comes from the study of « nose » whistlers (SMITH and HELLIWELL [22]) and it is found that at a distance of 5 Earth radii the electron density is approximately 100 cm^{-3} . WHIPPLE and FIREMAN [23] have determined an upper limit for the rate at which meteorites are etched by corpuscular radiation in interplanetary space. If the etching is done by protons with a speed of 1000 km s^{-1} , WHIPPLE [24] finds that the corresponding upper limit for the proton density is 40 cm^{-3} . The ion densities detected by Soviet space vehicles at a distance of 3-4 Earth radii have an upper limit of $(30 \div 60) \text{ cm}^{-3}$ (GRINGAUZ *et al.* [25]). Finally, BLACKWELL has stressed (private communication) that the accelerations observed in comet tails do not provide unambiguous evidence in support of the solar wind theory. The accelerations are a feature of only a fraction of the comets and might therefore be due to nothing more nor less than the occasional intense corpuscular emissions from the Sun which are already known to exist.

It seems therefore that until more is known about the Earth's interplanetary environment we must view with caution those ideas and theories that demand densities at the Earth's orbit in excess of 100 ions or electrons cm^{-3} . Furthermore, we must be prepared to accept the possibility that under quiescent conditions the actual density may be as much as an order of magnitude less.

In our discussions we shall adopt the solar wind theory, and assume that the particle density is typically 50 cm^{-3} , and that the velocity is approximately 500 km s^{-1} . If it turns out that there is no solar wind at all, then our remarks must be taken as applying only to those intermittent corpuscular emissions from the Sun for which there is definite evidence.

4. — The magnetosphere.

Close to the Earth the geomagnetic field is roughly that of a simple dipole. This has been known for a long time from the spherical harmonic analyses of the field components at the Earth's surface (CHAPMAN and BARTELS [26],

and more recently has been demonstrated in an impressive way by the study of whistlers, the discovery of the Van Allen belts (VAN ALLEN *et al.* [27]), and by the Argosy experiment (CHRISTOFILOS *et al.* [28]). This inner region of the geomagnetic field contains those flux-lines which intersect with the Earth's surface at latitudes below the auroral zones.

If R is the distance at which a flux-line crosses the geometric equatorial plane, and a is the radius of the Earth, then

$$(3) \quad \frac{R}{a} = \sec^2 \theta_0 ,$$

where θ_0 is the latitude at which the flux-line meets the Earth's surface. The auroral zones are situated at the geomagnetic latitudes of $\theta_0 = \pm 67^\circ$, approximately, and therefore from eq. (3) the flux-lines which fall into the auroral zones extend to a distance of $R = 6a$. Most likely this entire region bounded by the flux-lines extending to 6 Earth radii, rotates as a solid with the same angular velocity as the Earth itself. This conclusion follows from Ferraro's [29] law of isorotation and also that the magnetic stresses of $H^2/8\pi$ everywhere exceed the thermal, gravitational, and inertial forces. It is interesting to notice that the centrifugal force in the equatorial plane is equal to the Earth's gravitational attraction at a distance given by

$$(4) \quad \frac{R}{a} = \left(\frac{g_0}{a\omega^2} \right)^{\frac{1}{2}} ,$$

or, $R = 6.5a$, which is remarkably close to the limiting surface of the inner region of the geomagnetic field.

Perhaps the most outstanding feature of the inner region of the geomagnetic field is the presence of a geocorona consisting of the Van Allen radiation belts of energetic charged particles. There are two distinct radiation belts and their outer surfaces are situated at distances of roughly 2 and 4 Earth radii, respectively, from the Earth's centre. The inner belt contains a relatively large flux of high energy protons which are satisfactorily accounted for by the cosmic ray neutron albedo theory (SINGER [30]). The outer belt consists largely of energetic electrons (VERNOV *et al.* [31]) that are probably of solar origin; their average energy is about 25 keV and the flux density is $10^{11} \text{ cm}^{-2} \text{ s}^{-1}$. In the past it has been suggested that the outer radiation belt may be the immediate main source of the electron streams that produce the excitation phenomena and the X-rays (MEREDITH *et al.* [32]) in the auroral zones. However, the magnetic flux-lines which pass through the outer belt do not normally fall into the auroral zones, and if only for this reason, the suggestion as it stands cannot be accepted.

The outer or distant geomagnetic field contains those flux-lines which extend beyond 6 Earth radii in the equatorial plane, and which connect with the Earth's surface in the magnetic polar regions at latitudes above the auroral zones. Of this region very little is known directly, and yet the topology of the field in this region is of fundamental importance for understanding the physics of the Earth's immediate interplanetary environment. Although there is very little direct evidence for supporting the view, it is now widely agreed that the distant geomagnetic field is appreciably different from that of a simple dipole. Unfortunately, surveys of the distant geomagnetic field with magnetometers carried aboard the space probes have not been particularly revealing (SONNETT *et al.* [33]). Measurements made along a single trajectory cannot by themselves afford an overall picture of the time and space varying geometry of the field. It is possible that for this purpose other methods must be employed, such as the release of barium vapour which is photoionized by solar radiation and then is detected by resonance scattering as it maps out the magnetic field lines (HARRISON [33b]).

In 1896 BIRKELAND [34] suggested that magnetic storms and aurorae were caused by streams of fast electrons ejected from the Sun. STÖRMER developed the theory mathematically and studied the motion of a single charged particle in a geomagnetic dipole field which extended to infinity. The theory was subsequently criticized by SCHUSTER in 1911 [35] on the grounds that the Coulomb repulsion in the streams could not be neglected. LINDEMANN [36] suggested the possibility that the corpuscular streams consisted of neutral ionized gas; this suggestion was taken up and developed in the now classic work of CHAPMAN and FERRARO [37-39], CHAPMAN [40], FERRARO [41, 42].

CHAPMAN and FERRARO assumed that the ionized streams from the Sun had cross-sections which were large in comparison with the Earth and its immediate environment. Their major contribution was the study of a stream impacting against the geomagnetic field. The stream as it approached, swept up the distant geomagnetic field and compressed it, thus accounting for the abrupt increase in the horizontal component of the field at the Earth's surface, which is characteristic of the initial phase of a magnetic storm. In addition to compressing the geomagnetic field on the Sun's side of the Earth, the stream also flows round and pass the Earth leaving a hollow wake on the night side, in which the geomagnetic field is relatively undisturbed. CHAPMAN and FERRARO adopted the view that the hollow formed in the solar stream would persist to great distances. Furthermore, they considered the possibility of a ring-current about the Earth as a result of the polarization of the solar stream. MARTYN [43] attempted, on the basis of this idea, to developed an acceleration mechanism for injecting energetic particles into the auroral zones to account for auroral phenomena. These details, and other theories such as Alfvén's [44], are not considered in this discussion.

An approximate method for estimating the shape and dimensions of the hollow carved out in the solar stream is the following (*). The steady-state hydrodynamic equation is

$$(5) \quad \rho(\mathbf{V}\nabla)\mathbf{V} = \mathbf{j} \wedge \mathbf{B} - \nabla p.$$

For a hypersonic stream the thermal pressure can be neglected on the Sun's side of the Earth. Expressing $\mathbf{j} \wedge \mathbf{B}$ in terms of magnetic stresses it follows

$$(6) \quad \frac{1}{2}NMV^2 \simeq \frac{B^2}{8\pi},$$

as a first approximation, where N is the density of protons of mass M . Expressing the magnetic intensity in terms of B_0 at the Earth's surface:

$$B = B_0 \left(\frac{a}{R} \right)^3,$$

the corpuscular stream approaches the Earth to a distance of R given by

$$(7) \quad \frac{R}{a} \simeq \left(\frac{B^2}{4\pi NMV^2} \right)^{\frac{1}{2}}.$$

We now suppose that instead of a corpuscular beam we are dealing with a solar wind of $N = 50 \text{ cm}^{-3}$, and $V = 500 \text{ km s}^{-1}$. We then find from eq. (7) for $B_0 = 0.3$ gauss, that $R = 5.8a$, or slightly more, since the magnetic intensity will be higher than that assumed for a simple dipole. The first conclusion therefore is that even with a relatively mild solar wind there is no distant geomagnetic field beyond 6 Earth radii on the side of the Earth facing the Sun.

On the night side of the Earth one expects the hollow in the solar stream to ultimately close up because of the finite temperature (BEARD [45]) of the stream. Neglecting for the moment the effect of the magnetic field, it follows that the hollow should close up in a distance of the order

$$(8) \quad L = R \frac{V}{V_{\text{th}}},$$

where V_{th} is the thermal speed of the protons. For $V = 500 \text{ km s}^{-1}$, $R = 6a$, one finds L is 20 Earth radii for a temperature of $10^6 \text{ }^\circ\text{K}$, or 60 Earth radii

(*) This paper was prepared early in 1961 and a more detailed discussion on this subject has since been published by the author in *Geophys. Journ.*, 6, 479 (1962).

in the case of a temperature of 10^5 °K. The geomagnetic field at these distances will be quite weak, whatever form it might take.

Thus, as shown in Fig. 1, the hollow formed by the Earth in the interplanetary fluid has a tear-drop shape. Our calculations are exceedingly crude, but none the less they include, one hopes, the essential physical processes. The same highly simplifying assumptions have been used in more detailed analyses (PARKER [18], BEARD [45], ZHIGULEV and ROMISHEVSKII [46]) with essentially the same results. Progress in the direction of a more exact solution of this problem does not seem possible until more is known about the nature of the boundary layer formed where the solar wind is in contact with the geomagnetic field. Here one has perhaps the most formidable boundary layer problem known to physics. For example, PARKER [47] has shown that in all probability the outer geomagnetic surface exhibits Helmholtz instability in a manner similar to the surface of water which is in contact with moving air. Many authors have supposed that at the surface the electromagnetic fields vary abruptly according to the «superconductivity equation»

$$(9) \quad \nabla^2 \mathbf{X} - \lambda_s^{-2} \mathbf{X} = 0,$$

where $\lambda_s = (mc^2/4\pi ne^2)^{1/2} = 0.5 \cdot 10^5 N^{-1/2}$ is the «collision-free penetration depth». For a value of N equal to 50 cm^{-3} , this equation gives a boundary layer of only 1 km in thickness. It is doubtful whether such a treatment of the problem has any physical reality; where, for example, has the interplanetary medium gone to, which normally would have filled the hollow in the absence of the Earth? Such a situation can only obtain for periods of time less than a characteristic collision period for the surface currents at the boundary. Even if the mean free path is 1 A.U., the collision time is still only of the order of 1 day, and such a treatment is therefore possible only for transient phenomena, which in any case should be dealt with in terms of shock wave theory. With even very weak magnetic fields embedded in the solar wind one expects the flow past the Earth will be streamline, as for moderately compressible fluids flowing round a stationary obstacle.

The region about the Earth has been referred to as a hollow or cavity in the solar wind. This is misleading, for this region contains the geomagnetic field in which there is distributed a plasma of varying density. More appropriately, one might use Gold's [48] terminology and refer to it as the «magnetosphere». Its asymmetric shape affects only the distant geomagnetic field. On the Sun's side of the Earth the magnetic field is compressed and parted, and flux-tubes are carried round to the rear side of the Earth to form what we shall refer to as the «magnetic arch» (Fig. 1). The magnetic arch contains virtually all the magnetic flux which normally in a vacuum would extend symmetrically about the Earth to a distance greater than 6 Earth radii; it

therefore consists of the magnetic flux which comes from the polar regions. Hence, the magnetosphere contains two distinct regions: an inner region, which is symmetric about the Earth and contains the magnetic flux from latitudes below the auroral zones; and an outer region, which is mainly on the night side of the Earth, and contains the magnetic flux from latitudes above the auroral zones. The first region co-rotates with the Earth, and the second region is stationary.

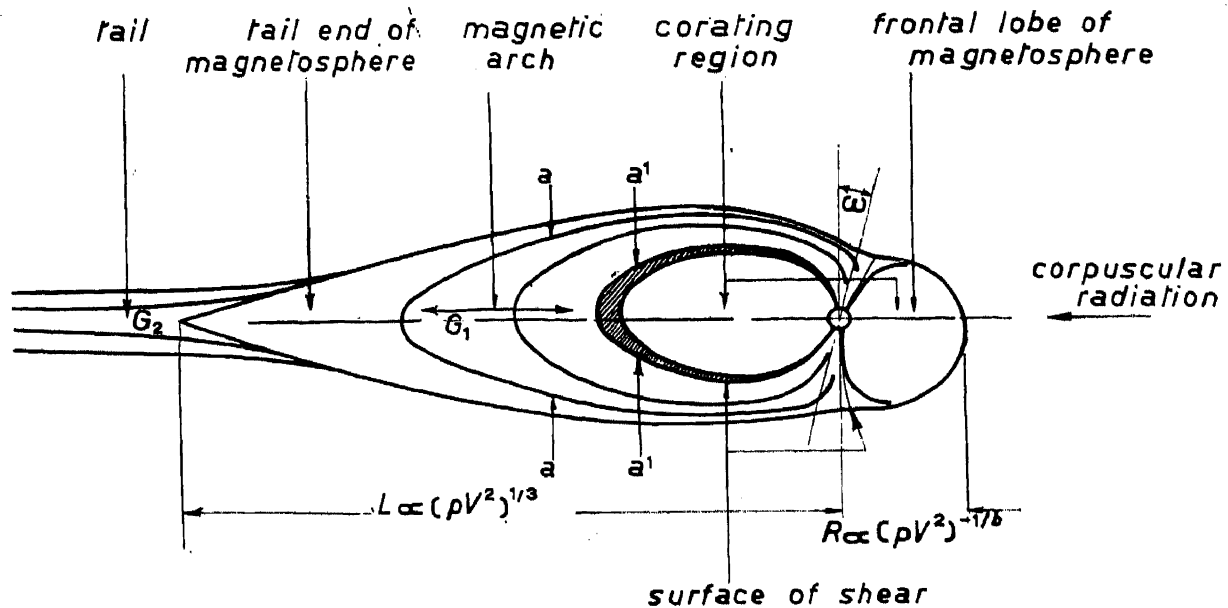


Fig. 1. — An asymmetric magnetosphere produced by corpuscular radiation from the Sun impacting against the Earth's magnetic field. The letters G_1 and G_2 indicate the possible locations of the gegenschein.

PARKER [18] has commented on the possibility that some of the polar magnetic flux might be captured or frozen into the solar wind, particularly under conditions of enhanced solar activity. Beyond the magnetic arch there is the strong possibility that the Earth possesses a magnetic tail directed away from the Sun. One would expect the magnetic tail to form at a distance of 20 or more Earth radii at the tail-end of the magnetosphere. JOHNSON [49] doubts whether a tail will form as it would consist of bundles of oppositely directed flux which would merge and cancel out each other by Sweet's mechanism (PARKER [50]). If flux is transported away from the Earth to form a tail, its ultimate dissipation must depend on collision processes, and it is therefore difficult to see why a tail should not form and extend to very considerable distances from the Earth.

Once we are reasonably sure of the geometry and properties of the magnetosphere, an outstanding problem will be the determination of its modes and periods of oscillation. Already the subject has received some attention (DUNGEY [51], OBAYASHI [52], KATO and WATANABE [53], JACOBS and

SINNO [54]). The micropulsations of the geomagnetic field show dissimilar characteristics for the day and night sides of the Earth. During daylight hours the fluctuations tend to be continuous (referred to as Pc's), but during the night they take the form of damped trains of oscillations (Pt's). On the night side the fluctuations of the magnetosphere are more complex, as one would expect, and we shall therefore consider briefly only the fluctuations occurring on the daylight side of the Earth.

In 1953 HOLMBERG [55] noticed that the micropulsations during daylight hours tended to fall into two distinct frequency wave-bands, having their peak values at 20 and 70 seconds per period, respectively. More recently, JACOBS and SINNO have analysed geomagnetic pulsation data recorded during the International Geophysical Year and find that Holmberg's original deductions apply to micropulsations observed at almost all latitudes below the auroral zones. The long period pulsations are synchronous over wide areas, and have maximum amplitudes approaching the polar regions amounting on occasions to a few tens of gammas ($\gamma = 10^{-5}$ gauss). The short period pulsations have maximum amplitudes of only a few gammas in moderate latitudes and are not synchronous over wide areas.

The correct approach to a problem of this kind is to set up the hydro-magnetic equations with the appropriate boundary conditions for the magnetosphere and study the modes of oscillation for arbitrary perturbations. A more elementary treatment is to use a string model, and by assuming that disturbances are propagated at the Alfvén phase velocity of $V = B/(4\pi\rho)^{1/2}$ along the magnetic field, to calculate the allowed frequencies at which the flux-lines oscillate. Obviously, such a model is grossly inaccurate, and can yield only the crudest results.

We assume that the flux-lines are anchored in the ionosphere and that the currents induced in this region are responsible for the field fluctuations at ground level. The flux-lines are symmetric about the geomagnetic equator and their perturbation by the impacting solar wind will excite, in addition to the fundamental frequency, only the odd overtones, as in the analogous case of a string fixed at both ends and disturbed symmetrically about its centre. The inclination of the magnetic equatorial plane to the ecliptic must mean that there are also even overtones of high order, but as their amplitude is small they can be neglected.

As a first approximation we suppose that the allowed frequencies are harmonic and given by

$$(10) \quad \nu_n = \frac{1}{2} n \left(\int V^{-1} ds \right)^{-1},$$

as for an ordinary string, where $n = 1, 3, 5, \dots$, and the integral is along a flux-line between the conjugate nodal points in the ionosphere. The Alfvén

phase velocity is not greatly sensitive to variations in density, and assuming that ρ is constant above, say, 1000 km, the corresponding eigenperiods to eq. (10) are

$$(11) \quad T_n \simeq 4 \left(\frac{R}{a} \right)^4 \frac{a}{nB_0} (\pi\rho)^{\frac{1}{2}}.$$

For $R/a=5$, and a number density of $N=100$ protons cm^{-3} , it follows that $T_1=100$ s and $T_3=33$ s.

On the basis of this model discontinuities in displacement are not expected and therefore the amplitudes of the Fourier components converge at least as rapidly as n^{-2} , and the amplitude of v_3 will therefore be an order of magnitude or more less than that of v_1 , in agreement with observation. In spite of the limitations of the string model these results are close enough to Holmberg's periods of 70 and 20 s for the long and short wave-bands to suggest that micropulsations are the free oscillations of the magnetosphere.

Pursuing the idea of a string model even further, we know that if the phase velocity is not constant along a flux-line, the allowed frequencies in general are not harmonic (the overtone frequencies are not integral multiples of the fundamental). The geomagnetic field strength varies as r^{-3} , and the phase velocity must therefore alter in a similar manner. Consider a length l along ρ a flux-line with its end points fixed rigidly and symmetrically about the equatorial plane. Using perturbation theory and assuming that the field varies as the cube power along the line, then for $l \ll R$

$$(12) \quad v'_n = \frac{1}{2} n \left(\int V^{-1} ds \right)^{-1} \left\{ 1 + \frac{3l}{2R} \left(1 - \frac{4}{\pi^2 n^2} \right) \right\}^{\frac{1}{2}}.$$

This equation illustrates the general rule that if the phase velocity decreases towards the mid-point of the string the overtones of frequency v'_n are less than nv'_1 . If the result (12) is extrapolated to the limit of $l \simeq 2R$, which strictly is not a valid procedure, then

$$(13) \quad v'_n \simeq 2v_n \left(1 - \frac{3}{\pi^2 n^2} \right)^{\frac{1}{2}},$$

and therefore T'_1 is slightly less than T_1 obtained from eq. (1). The ratio of the periods of the fundamental and the second overtone is $T'_1/T'_3 \simeq 3.5$. The fact that Holmberg's ratio of $T'_1/T'_3 = 70/30$ is also 3.5 is perhaps a fortuitous coincidence in view of the impermissible extrapolation of a perturbation treatment. However, on the basis of the string model one expects in general that $T'_1/T'_3 > 3$, in agreement with observation. Failing a more detailed knowledge of the magnetosphere, the string model in spite of its crudity is physically able to explain at least qualitatively some of the properties of micropulsations.

5. – Origin of auroral streams.

In spite of the sustained interest the aurorae have excited in living memory, we still do not understand the basic mechanisms responsible for this remarkable phenomenon (HARANG [20], FERRARO [56], CHAMBERLAIN and MEINEL [57], STORMER [58], ELVEY [59], PARKER [60]). We shall leave aside most of the problems connected with the aurorae, and concentrate on only those which appear to be specifically related to the general nature of the Earth's immediate environment.

Since the observation by MEINEL of Doppler-shifted hydrogen lines in low latitude homogeneous arcs there can remain no doubt that solar corpuscular emissions play a fundamentally important part in the excitation of auroral displays. There is evidence that the displays are excited by both electron and proton streams occurring either simultaneously or separately in different types of auroral structure. Furthermore, one of the most revealing features of the auroral displays comes from the fact that they are essentially a night time phenomenon, as shown by radio studies (see for example CURRIE *et al.* [61]). The excitation of the aurorae by corpuscular streams from the Sun and their occurrence on the night side of the Earth can only be reconciled in the framework of a terrestrial environment which possesses marked asymmetric properties. The acceleration mechanisms necessary for producing energetic particles which can penetrate to the 100 km level must also find their natural place in such an asymmetric environment.

We have seen that the distant geomagnetic field is distributed predominantly on the night side of the Earth, and it is possible to attribute to this fact alone many of the singular properties of the aurorae. We shall suggest three reasons why the auroral displays occur on the night side of the Earth.

1) Already it has been pointed out that the magnetosphere consists of a region which co-rotates with the Earth and a more distant, highly asymmetric region which is stationary. There must therefore be an intermediate region or surface of shear, as shown in Fig. 1. Intuitively one expects this region to consist of highly agitated plasma, and disordered magnetic flux, within which Fermi interactions (PARKER [18]) are the efficient mechanism for accelerating protons. It is significant that this region is in the form of a shell which envelops the night side of the Earth and extends down into the auroral zones close to the Earth's surface.

2) As the Earth rotates the flux-lines within the magnetic arch must change their relative position, as pointed out by JOHNSON [49]. Thus the flux-line marked « α » in Fig. 1, becomes the flux-line marked « α' » when the Earth has rotated through 180° about its axis.

Consider a tube of flux containing protons that have leaked in from the solar wind. Their longitudinal velocity $V_{\parallel}$ is of the order of 500 km s^{-1} . Their transverse drift velocity in the equatorial plane is

$$V_D = V_{\parallel}^2 \frac{Mc}{Be\varrho_R},$$

for $V_{\parallel} \ll V_{\perp}$, where ϱ_R is the radius of curvature of the flux-line. Assuming that $B \sim 10^{-2}$ gauss, and $\varrho_R \sim 5a$, it follows that $V_D \sim 10^4 \text{ cm s}^{-1}$, which is very much less than the speed of the flux-tubes. As a first approximation it can be taken that the protons are confined to a given flux-tube. We can therefore use the longitudinal adiabatic invariant

$$J = \int V_{\parallel} dl,$$

in the sense that the integral is along a given tube of force, and is evaluated between the mirror points. Thus, if L is the length of the magnetic tube of force at any instant, then $V_{\parallel} \propto L^{-1}$. If the tube changes its length by as much as a factor of 10, then the original energy of approximately 1 keV at « a » becomes 100 keV at « a' », which is sufficient for penetration to auroral altitudes. Hence, the protons are collected on the Sun's side of the Earth and are virtually squeezed out of the tubes of force on the night side of the Earth. Practically all particles in the magnetic arch are in a state of cyclic acceleration and deceleration, but the maximum periodic change in energy is associated with those flux-lines that change their length by the greatest amount. It is evident that these particular flux-lines fall into the auroral zones. Also, when they are at the position « a » they are adjacent to the solar stream, and are better situated for being charged with energetic protons than the flux-lines buried in the interior of the magnetic arch. This mechanism, together with the previous proposal, has the virtue that it explains not only why the aurorae are situated on the night side, but also why they occur in relatively narrow zones.

3) Both the above proposals offer efficient mechanisms for accelerating protons, but are unable to account for energetic electrons. The energy required by the electrons is not large (FAN [62]) and elsewhere it has been suggested (HARRISON [63]) that collective electrostatic interactions between the protons and electrons will cause the electrons to be dragged down with the proton streams and thus gain transverse energy in the converging magnetic flux-lines. Alternatively, as PIDDINGTON [64] has pointed out, the polar flux-lines are possibly twisted about each other in a helical fashion. This would certainly occur to some degree if any resistance is offered to the motion of the

plasma in the magnetic arch. This plasma is churned around and there must therefore be some impedance to the motion of the magnetic flux. The twisting of the polar bundles of magnetic flux into helical-like paths must mean that electric currents flow from the magnetic arch down into the magnetic polar regions, radially outwards to the auroral zones, and then up along the flux-lines back into the magnetic arch. Or in other words, there is an electric field which drags electrons down into the auroral zones. HINES [65] has discussed similar phenomena in the polar regions although apparently he adopts the point of view that the currents are located in regions close to the Earth's surface. It is clear that the rotational deformation or twist of the polar magnetic flux will establish inwardly and outwardly directed electric fields which will contribute to the auroral and polar phenomena, and may well produce the electron streams which appear to be a substantial feature of auroral displays, particularly in their latter stages.

6. — The Earth's gas-magnetic tail.

HOUZEAU [66] and EVERSHED [67, 68] have independently proposed the theory that the Earth possesses a gaseous tail directed away from the Sun, very much like the tail of a comet. They were led to this conclusion in their attempts to account for the gegenschein. Owing to the difficulties of observing the gegenschein very little appears to be known about it (MITRA [69]). BARNARD [70-74] recorded his visual observations over a number of years and refers to BRORSSEN [75] as the original discoverer, followed later by BACKHOUSE [76]. However, it appears that Humboldt also knew of the gegenschein at the end of the eighteenth century, and possibly there were others before him. More recently the gegenschein has attracted the attention of a number of Russian workers (see HOPE [77]) who have revived the gaseous tail theory.

The recorded observations of the gegenschein are not all in agreement, and one must regard the little that is known with a certain amount of reservation. From the visual observations of BARNARD, EVERSHED, ASTAPOVITCH [78] and FESENKOV [79] we learn that the gegenschein is a faint luminosity of the night sky, close to the ecliptic, and almost directly opposite the Sun. It usually appears elliptical in shape with the major axis lying nearly along the ecliptic. Because of its faintness it has been found difficult to assign any accurate dimensions to the gegenschein; generally, it extends about 10° along the major axis and 6° along the minor axis, although on other occasions it may attain a size as large as 40° . With the exception of BARNARD, most observers agree that the centre of the gegeneschein is displaced to the west of the counter-solar point by about 3° . Another characteristics feature is that it frequently appears «spotty». DIVARI [80], BEHR and SIEDENTOPF [8] and

ROACH and REES [81] have made photometric studies of the gegenschein, and the latter believe there is evidence for thinking that the gegenschein and the zodiacal light have their origin in separate and distinct concentrations of particles. FESENKOV [79] and KARIMOV [82], on the basis of spectroscopic studies, have claimed that the gegenschein has an emission spectrum similar to that of the night air glow; it now seems (PARIISKII and GINDILIS [83]) that they were mistaken and the spectrum is continuous and resembles that of the zodiacal light with the exception of some minor features.

The gaseous tail theory of the origin of the gegenschein encounters the hitherto unsolved problem of accounting for a non-divergent jet of gas projecting away from the night-side of the Earth. PIDDINGTON [84] has made the remarkable suggestion that the gegenschein is located within the Earth's magnetic tail. PIDDINGTON proposes that the magnetic tail is an occasional appendage which may persist as a remnant of a magnetic storm, and that as the distant magnetic field recovers some of its symmetry, the energetic particles in the tail region are pulled in to populate the outer Van Allen belt. A different point of view has been adopted in this discussion; even the most tenuous solar wind is adequate to maintain a complete asymmetry of the distant geomagnetic field, and because the field intensity falls off rapidly as one moves away from the Earth, the incidence of geomagnetic storms will not change qualitatively the general tear-drop shape of the magnetosphere with its trailing filamenting tail of magnetic flux.

Soviet observations by ASTAPOVITCH, DIVARI, FESENKOV, KARIMOV, and ROZHKOVSKI, both visual and photometric, have established that there is a parallax shift of the gegenschein as the Earth revolves. They calculate that the gegenschein is at a mean distance of 20 Earth radii. This distance would place the luminosity at the tail end of the magnetic arch. Whatever the nature of the particles responsible for the gegenschein luminosity, whether they are of atmospheric origin or not, one can account for the elliptical shape of the luminosity by supposing that they are concentrated at the tail-end of the magnetic arch in the region marked G_1 , in Fig. 1 (HARRISON [33b]).

However, although this theory is attractive, one hesitates to accept outright the results of the parallax measurements. Accurate measurements to within $\pm 0.5^\circ$ even of the position of the axis of the zodiacal light are very difficult to make. An error of as little as $\pm 1^\circ$ in the parallax measurements would make the position of the gegenschein uncertain as to whether it is 15 or 30 Earth radii distant. Such an error is not improbable. FESENKOV [85] claims that the deviation of the axis of the zodiacal light from the plane of the ecliptic is completely negligible, although other workers and more recently BLACKWELL [21] have shown that the planes of the zodiacal light and the ecliptic are by no means coincident. In view of the variable nature of the gegenschein, and that sometimes pronounced changes occur in even a single

night, it would seem that the problem of determining the parallactic shift of the gegenschein is at least as difficult as determining the axis of the zodiacal light. Until further observations are made, and there is a measure agreement between them, we cannot claim to know exactly where the gegenschein is situated. It may be within the magnetic arch, as previously suggested, and this is theoretically plausible, or it may be situated in the magnetic tail in the way PIDDINGTON has suggested, as for example, at the points G_2 in Fig. 1, and can therefore extend for a considerable distance from the Earth. Independently of where exactly the gegenschein is located, it is interesting to notice that the early gaseous tail theories are now made more plausible with the concept of a magnetized medium, and elsewhere the author (HARRISON [86]) has proposed the name of «gas-magnetic tail» for the region in which the gegenschein is situated. All observers agree that the gegenschein is «spotty» or granular in appearance; this is exactly what one would expect from a gaseous, ionized medium in which there are embedded ropes of magnetic flux.

Let us assume that the solar wind is directed radially away from the Sun. Since the Earth has an orbital velocity of 30 km s^{-1} , it is clear that the major axis of the magnetosphere will be inclined in the plane of the ecliptic by an amount α to a line joining Earth and Sun. The angle α is given by the relation

$$(15) \quad v_r = v_e \operatorname{ctg} \alpha,$$

where v_r is the velocity of the solar wind, and $v_e = 30 \text{ km s}^{-1}$ is the Earth's orbital velocity. If now the gegenschein is situated in the magnetosphere, or in the trailing filaments of its magnetic tail, then it will be seen to be displaced to the west of the counter-solar point. This is exactly what is observed, and most observers have determined the value of α as 3° . Putting this values of α in eq. (15) yields a solar wind velocity of 600 km s^{-1} (HARRISON [86]). It is remarkable how the observed amount of the western displacement unites together the various tentative assumptions that have been made in the course of this discussion. A physical picture consisting of a solar wind of approximately 500 km s^{-1} , which interacts with the distant geomagnetic field to produce an elongated magnetosphere that is responsible for all the phenomena distributed in an asymmetric way about the Earth, such as the aurorae and the gegenschein, would not be possible if there were not a western displacement of the gegenschein of about 3° .

At the present stage of our knowledge of the Earth's immediate interplanetary environment we cannot escape the conclusion that our ideas are essentially highly speculative and are founded on the most meagre of observations. Reliable and unambiguous information is still pitifully small. An approach to the subject in the proper scientific manner compels us to confess that we know exceedingly little in this branch of physics; we do not know

whether there is a solar wind, which the present author and many others have rather rashly used as the basis of their theories; certainly we do not know the nature of the Earth's distant magnetic field, and even less do we know whether there exists a « magnetic arch » or a « gas-magnetic tail » far above us in the night sky.

REFERENCES

- [1] R. C. CARRINGTON: *Monthly Notices Royal Astron. Soc.*, **20**, 13 (1859).
- [2] C. CHREE and J. M. STAGG: *Phil. Trans.*, **227**, 21 (1927).
- [3] W. M. H. GREAVES and H. W. NEWTON: *Monthly Notices Royal Astron. Soc.*, **88**, 556 (1928).
- [4] W. M. H. GREAVES and H. W. NEWTON: *Monthly Notices Royal Astron. Soc.*, **89**, 84, 641 (1929).
- [5] H. W. NEWTON and W. JACKSON: *7-th Report of Commission for Study of Solar Terrestrial Relations* (Paris, 1951).
- [6] A. B. MEINEL: *Astrophys. Journ.*, **113**, 50, 583 (1951).
- [7] L. R. O. STOREY: *Phil. Trans.*, A **246**, 113 (1954).
- [8] A. BEHR and H. SIEDENTOPF: *Zeits. f. Astrophys.*, **32**, 19 (1953).
- [9] S. CHAPMAN: *Smithson. Contrib. Astrophys.*, **2**, 1 (1957).
- [10] S. CHAPMAN: *Threshold of Space*, ed by M. ZELIKOFF (London, 1958).
- [11] L. BIERMANN: *Zeits. f. Astrophys.*, **25**, 161 (1948).
- [12] L. BIERMANN: *Zeits. f. Astrophys.*, **29**, 274 (1951).
- [13] L. BIERMANN: *Zeits. f. Naturforsch.*, **7 a**, 127 (1952).
- [14] L. BIERMANN: *Observatory*, **77**, 109 (1957).
- [15] N. T. BOBROVNIKOV: *Publ. Lick. Obs.*, **17**, P. 2 (1931).
- [16] K. WURM: *Mitt. der Sternw. Hamburg-Bergedorf*, **8**, 51 (1943).
- [17] K. O. KIEPENHEUER: *The Sun*, ed by G. P. KUYPER (Chicago, 1953).
- [18] N. E. PARKER: *Astrophys. Journ.*, **129**, 664 (1958); *Phys. of Fluids*, **1**, 171 (1958).
- [19] E. N. PARKER: *Astrophys. Journ.*, **132**, 175 (1960).
- [20] L. HARANG: *The Aurorae* (New York, 1951).
- [21] D. E. BLACKWELL and M. F. INGHAM: *Monthly Notices Royal Astron. Soc.*, **122**, 113, 129, 143 (1961).
- [22] R. L. SMITH and R. A. HELLIWELL: *Journ. Geophys. Res.*, **65**, 2583 (1960).
- [23] F. L. WHIPPLE and E. L. FIREMAN: *Nature*, **183**, 1315 (1959).
- [24] F. L. WHIPPLE: *Journ. Geophys. Res.*, **64**, 1635 (1959).
- [25] K. I. GRINGAUZ, V. G. KURT, V. I. MOROZ and I. S. SHKLOVSKII: *Soviet Astron.*, **3**, 680 (1960).
- [26] S. CHAPMAN and J. BARTELS: *Geomagnetism* (Oxford, 1940).
- [27] J. A. VAN ALLEN, C. E. MCILVAIN and G. H. LUDWIG: *Journ. Geophys. Res.*, **64**, 271 (1959).
- [28] N. C. CHRISTOFILOS: *Journ. Geophys. Res.*, **64**, 869, (1959).
- [29] V. C. A. FERRARO: *Month. Not. Roy. Astron. Soc.*, **97**, 458 (1957).
- [30] S. F. SINGER: *10th Int. Astronautical Fed. Conf.* (1959).
- [31] S. N. VERNOV: *Dokl. Akad. Nauk S.S.S.R.*, **125**, 304 (1959).
- [32] L. H. MEREDITH, H. B. GOTTLIEB and J. A. VAN ALLEN: *Phys. Rev.*, **97**, 201 (1955).

- [33] C. P. SONNETT, D. L. JUDGE, A. R. SIMS and J. H. KELSO: *Symposium on Physical Processes in the Sun-Earth Environment* (Ottawa, 1959), p. 67.
- [32b] E. R. HARRISON: *Geophys. Journ.*, **6**, 462 (1962).
- [34] K. BIRKELAND: *Arch. Sci. Phys. Genève*, **4**, 497 (1896).
- [35] A. SCHUSTER: *Proc. Roy. Soc.*, **A85**, 45 (1911).
- [36] F. A. LINDEMANN: *Phil. Mag.*, **38**, 669 (1919).
- [37] S. CHAPMAN and V. C. A. FERRARO: *Terr. Mag. and Atmos. Elec.*, **36**, 77, 171 (1931).
- [38] S. CHAPMAN and V. C. A. FERRARO: *Terr. Mag. and Atmos. Elec.*, **37**, 147, 421 (1932).
- [39] S. CHAPMAN and V. C. A. FERRARO: *Terr. Mag. and Atmos. Elec.*, **38**, 175 (1933).
- [40] S. CHAPMAN: *Rev. Mod. Phys.*, **32**, 919 (1960).
- [41] V. C. A. FERRARO: *Journ. Geophys. Res.*, **57**, 15 (1952).
- [42] V. C. A. FERRARO: *Rev. Mod. Phys.*, **32**, 934 (1960); *Journ. Geophys. Res.*, **65**, 3951 (1960).
- [43] D. F. MARTYN: *Nature*, **167**, 92 (1951).
- [44] H. ALFVÉN: *Cosmical Electrodynamics* (London, 1950).
- [45] D. B. BEARD: *Journ. Geophys. Res.*, **65**, 3559 (1960).
- [46] V. N. ZHIGULEV and E. A. ROMISHEVSKII: *Sov. Phys. Doklady*, **4**, 859 (1960).
- [47] E. N. PARKER: *Astrophys. Journ.*, **128**, 664 (1958).
- [48] T. GOLD: *Journ. Geophys. Res.*, **64**, 1219 (1959).
- [49] F. S. JOHNSON: *Journ. Geophys. Res.*, **65**, 3049 (1960).
- [50] E. N. PARKER: *Journ. Geophys. Res.*, **62**, 509 (1957).
- [51] J. W. DUNGEY: *Physics of the Ionosphere* (1954), p. 229.
- [52] T. ODAYASHI: *Ann. d. Geophys.*, **14**, 464 (1958); *Rep. Ionosph. Res. Jap.*, **12**, 301 (1958).
- [53] Y. KATO and T. WATANABE: *Journ. Geophys. Res.*, **63**, 741 (1958).
- [54] J. J. JACOBS and K. SINNO: *Geophys. Journ.*, **3**, 333 (1960).
- [55] E. R. R. HOLMBERG: *Roy. Astron. Soc. Geophys. Suppl.*, **6**, 467 (1953).
- [56] V. C. A. FERRARO: *Adv. Phys.*, **2**, 265 (1953).
- [57] J. W. CHAMBERLAIN and A. B. MEINEL: *The Earth as a Planet*, ed. by G. P. KUIPER (Chicago, 1954).
- [58] C. STÖRMER: *The Polar Aurorae* (Oxford, 1955).
- [59] C. T. ELVEY: *Proc. U. S. Nat. Acad. Sci.*, **43**, 63 (1957).
- [60] E. N. PARKER: *Proc. I.R.E.*, **47**, 239 (1959).
- [61] B. W. CURRIE, P. A. FORSYTH and F. F. VAWTER: *Journ. Geophys. Res.*, **58**, 179 (1953).
- [62] C. Y. FAN: *Astrophys. Journ.*, **119**, 294 (1954).
- [63] E. R. HARRISON: *Nature*, **187**, 383 (1960).
- [64] J. H. PIDDINGTON: *Geophys. Journ.*, **2**, 173 (1959).
- [65] C. O. HINES: *Journ. Geophys. Res.*, **65**, 141 (1960).
- [66] J. HOUZEAU: *Ciel et Terre*, **1**, 517 (1888).
- [67] J. EVERSIED: *Observatory*, **19**, 57 (1899).
- [68] J. EVERSIED: *Observatory*, **22**, 57, 272 (1902).
- [69] S. K. MITRA: *The Upper Atmosphere*, 2nd ed. (Calcutta, 1952).
- [70] E. E. BARNARD: *Astron. Journ.*, **7**, 186 (1888).
- [71] E. E. BARNARD: *Astron. Journ.*, **11**, 19 (1892).
- [72] E. E. BARNARD: *Astron. Journ.*, **13**, 169 (1893).
- [73] E. E. BARNARD: *Astron. Journ.*, **17**, 151 (1897).
- [74] E. E. BARNARD: *Pop. Astron.*, **7**, 169 (1899).

- [75] T. J. C. A. BRORSEN: *Astr. N.*, 998 (1855).
- [76] T. W. BACKHOUSE: *Monthly Notices Royal Astron. Soc.*, **36**, 46 (1875).
- [77] E. R. HOPE: *The Earth's Exterior Atmosphere and the Counter glow*, *Scientific Inf. Service*, Report T65A (Defence Research Board of Canada, 1957), contains translations of a number of Russian papers on the gegenschein.
- [78] I. S. ASTAPOVITCH: *Priroda*, **1**, 25 (1950).
- [79] V. G. FESENKOV: *Astron. Žurn.*, **27**, 89 (1950).
- [80] N. B. DIVARI: *Astron. Žurn.*, **26**, 355 (1949).
- [81] F. E. ROACH and M. H. REES: *The Airglow and Aurorae*, ed by E. B. ARMSTRONG and A. DALGARNO (London, 1956).
- [82] M. G. KARIMOV: *Astron. Zurn.*, **29**, 633 (1952).
- [83] N. N. PARIISKII and L. M. GINDILIS: *Soviet Astron.*, **3**, 992 (1960).
- [84] J. H. PIDDINGTON: *Journ. Geophys. Res.*, **65**, 93 (1960).
- [85] V. G. FESENKOV: *Sov. Astron.*, **2**, 299 (1958).
- [86] E. R. HARRISON: *Nature*, **189**, 993 (1961).

The Production of High-Energy Particles in the Sun and in the Solar System.

E.-Å. BRUNBERG

The Royal Institute of Technology - Stockholm

1. - Introduction.

Let me first start this lecture by pointing out that very little is known about the mechanism, which produces particles of high energies on the sun, although theories have been put forward to explain the experimental fact that during large solar flares, particles are accelerated to high energies which are observed on the earth. In this connection we shall also discuss different theories for the propagation of flare particles from the sun to the earth.

Whilst all agree to the fact that high energy particles are produced during solar flares, there is no such general understanding about the origin of cosmic radiation. The energy spectrum of this radiation extends from about 10^9 eV to 10^{18} or 10^{19} eV. Different authors claim that radiation is generated in intergalactic space (BIERMANN, HEIDMANN) or in interstellar space (Fermi and others) or at the supernovae (BAADE, SHKLOWSKI and others). TELLER and RICHTMEYER have suggested that most of the radiation is of local origin, *i.e.* generated within our solar system. ALFVÉN has developed a theory along these lines and at the end of the lecture the basic ideas of his theory will be explained.

Before we end this introduction, let me show the following figure (Fig. 1), which has been given by Prof. SIMPSON and which shows the spectra of energetic particles observed in the solar system.

a) Is the spectrum of the cosmic radiation, whose origin is still unknown although, as mentioned above several theories have been suggested to explain its origin.

b) Is the first spectrum of protons, derived from a solar flare (February 1956).

c) Is also from a flare, but not with so many high energy particles as in b).

d) and *e)* represent the spectrum of protons and electrons in the radiation belts around the earth. This particular radiation will be discussed in another lecture.

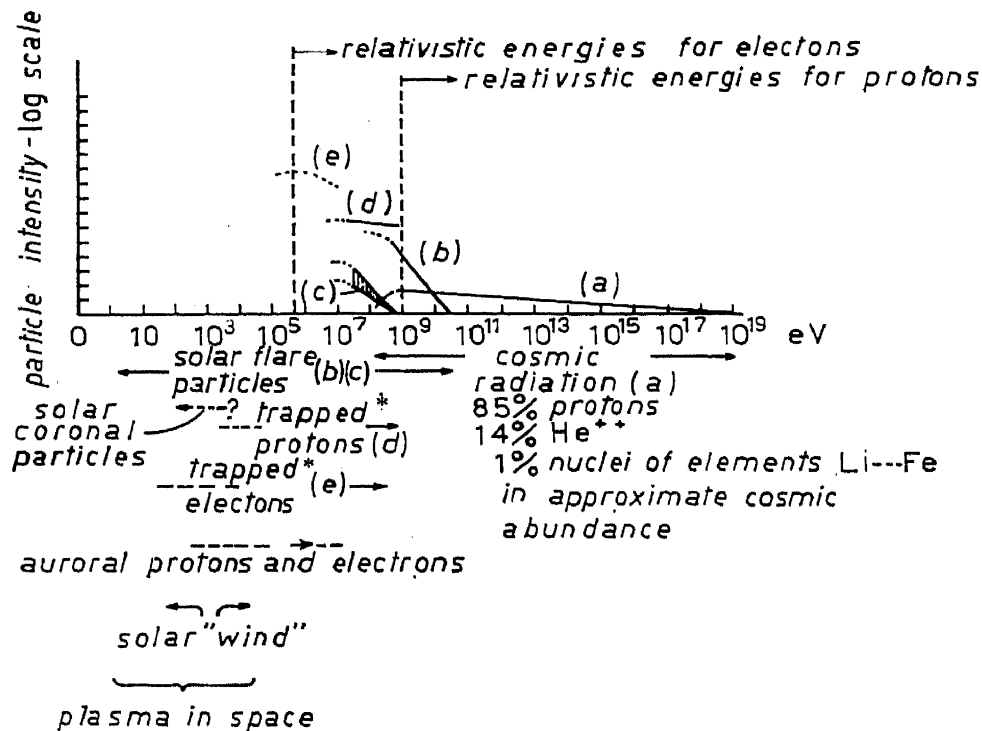


Fig. 1. - - - - Indicates unknown extension of spectrum. * Trapped in geomagnetic field, some of these particles are from neutron decay.

2. - What is a flare?

First a short review of some well-known facts about the sun. The sun is a sphere of hot ionized gases, mainly hydrogen. The solar energy is produced by nuclear reactions in a region around the center of the sun, where a temperature of the order of $20 \cdot 10^6$ degrees is probable.

There are three different layers in the outer part of the sun: the photosphere, the chromosphere and the corona.

The *photosphere* is the source of the visual radiation from the sun. It has a radius of $6.95 \cdot 10^8$ m and a thickness of 10^6 m. The radiation has a continuous spectrum with the Fraunhofer absorption lines.

The *chromosphere* has a rapidly changing structure and gives a spectrum of emission lines. Its height is about 10^7 m.

The *corona*, finally, has a continuous spectrum, due to scattered light from the photosphere. It is highly ionized with a temperature of about 10^6 degrees and extends sometimes out to ten solar radii.

Sunspots are regions in the photosphere with lower temperature than their environment. According to ALFVÉN the sunspots originate probably from the inner parts of the sun and are propagated to the surface by magneto-hydrodynamic processes. Spots usually occur in pairs and are associated with strong

magnetic fields (of the order of several hundred gauss, but up to 4500 gauss has been measured).

Within the chromosphere and in the vicinity of sunspots there sometimes occurs a sudden increase in optical as well as radio emission, which during large events are both easily observed in spite of other normal emissions. This phenomenon is called a *flare*. The flare appears as a luminous region just above the photosphere. The flare intensity has its maximum after a few minutes and then gradually decreases during a time of the order of hours. The total energy, spent in a large flare, has been calculated to be between $(10^{30} \div 10^{33})$ erg. It is supposed that this energy has been stored in the magnetic field in the vicinity of a sunspot. The mechanism, which triggers the flare energy so suddenly, is not known. It has, however, been suggested by GOLD and HOYLE that when special magnetic field configurations occur, the magnetic forces act so as to push plasma away from the sunspot region and magnetic energy is converted into kinetic energy and heat. The energy which is used for the acceleration of plasma thermal energies up to relativistic values of cosmic ray energies has been estimated to be only a small fraction of the total flare energy.

Solar flares are divided in different classes, depending on their surface area: 1^- , 1, 2, 3 and 3^+ where 1^- mean a flare with a surface $< 100 \cdot 10^{-6}$ of the hemisphere and 3^+ means a flare with a surface $> 1200 \cdot 10^{-6}$ of the hemisphere.

3. - Flare produced particles and their propagation between the sun and the earth.

Acceleration of particles to relativistic energies in a conducting medium, as in the highly ionized regions of the sun, needs a very special mechanism. A change in the magnetic flux with time induces currents within the conducting medium, *i.e.* a betatron effect, but the accelerator or «dynamo» is short circuited and gives strong currents with low particle energies instead of a low current, *i.e.* few particles of high energy. Furthermore it is difficult for the accelerated particles to escape from the general solar magnetic field.

PARKER has suggested that some form of Fermi's acceleration mechanism for cosmic ray particles in interstellar space might be applicable also as a flare mechanism for particle acceleration. Let us, however, first review the basic idea of Fermi's acceleration mechanism.

FERMI assumes that in interstellar space plasma clouds with «frozen in» magnetic fields are moving at random with an average velocity of say $v = 30$ km/s. If charged particles, say protons, approach such a magnetic field they are deflected. The particles in other words collide with the clouds, which have enormous masses, compared to the particle masses. In a collision a particle gains or loses energy, depending on whether the velocity vectors u and v of the particle and the cloud before collision are pointing in opposite directions or in the same direction. u is the averaged velocity of the particles.

It can be shown that the gain in energy is larger than the loss and that the net average gain per collision to as order of magnitude is given by

$$\delta W = \frac{v^2}{c^2} \cdot mc^2 ,$$

where $mc^2 = W$ = particle energy inclusive rest energy. After dn collisions the energy is

$$\frac{dW}{W} = \frac{v^2}{c^2} \cdot dn ,$$

$$\ln W = \frac{v^2}{c^2} \cdot n + \text{const} ,$$

$$W = m_0 c^2 \cdot \exp \left[\frac{v^2}{c^2} \cdot n \right] .$$

If $v/c = 10^{-4}$, $\delta W = 10^{-8} \cdot 900 \cdot 10^6 \text{ eV} \approx 10 \text{ eV/collision}$ in the non relativistic range. This also means we need $\approx 10^8$ collisions for an energy gain by a factor e if losses are neglected.

PARKER applies this idea to the flare process. He assumes that the magnetic field of the flare becomes unstable, part of the magnetic energy is released and results in magneto-hydrodynamic waves in the medium. A charged particle will spiral in the field along the lines of force for a certain distance and then be reflected as in Fermi's model from a magnetic inhomogeneity which moves with the Alfvén velocity.

Taking into account energy losses due to electrostatic interactions between the accelerated particles and the ions and electrons of the surrounding plasma, Parker states that an acceleration is possible. He finds that a proton may reach 10^9 eV in a few minutes, when assuming a magnetic field of 500 gauss, and he is also able to calculate a spectrum, which is not in disagreement with that observed at the earth during strong flares.

In order to check an acceleration mechanism one must know the spectrum of the produced flare particles at the source. This means that we must have an idea of how this spectrum is changed when the particles reach the earth, where we measure the spectrum and consequently we have to study the propagation mechanism between the sun and the earth.

We shall now discuss different models for the propagation of solar flare high energy particles between sun and earth. One of the first attempts to solve this problem was made by FIROR. Obviously the earth's magnetic field deflects particles when they approach the earth. Depending on their momentum and position far away in space, where the particles start their motion towards the earth, they hit the earth within certain regions or impact zones. Between the impact zones there are forbidden regions. Using electronic computers Firor calculated impact zones on the earth for particles of different momenta, starting from the sun and being deflected by the earth's dipole field.

He assumes, however, that there is a field free region between the sun and the earth. This is certainly not true. The existence of a general solar magnetic field must be taken into account, which will modify the Firor impact zones.

Before discussing theories of propagation we may sum up the observational facts as regards cosmic ray intensity in connection with a flare outburst. During a period of 23 years, 6 really intense flares have been observed followed by a large increase in cosmic ray intensity at the surface of the earth. The most intense so far was the one of February 23, 1956. The neutron recording instruments at high latitudes measured increases of $6 \div 7$ times the pre-flare value.

It is, however, important to remember that there are many geophysical phenomena connected with a flare and the next series of diagrams, which is put together by K. ANDERSON, gives a good survey:

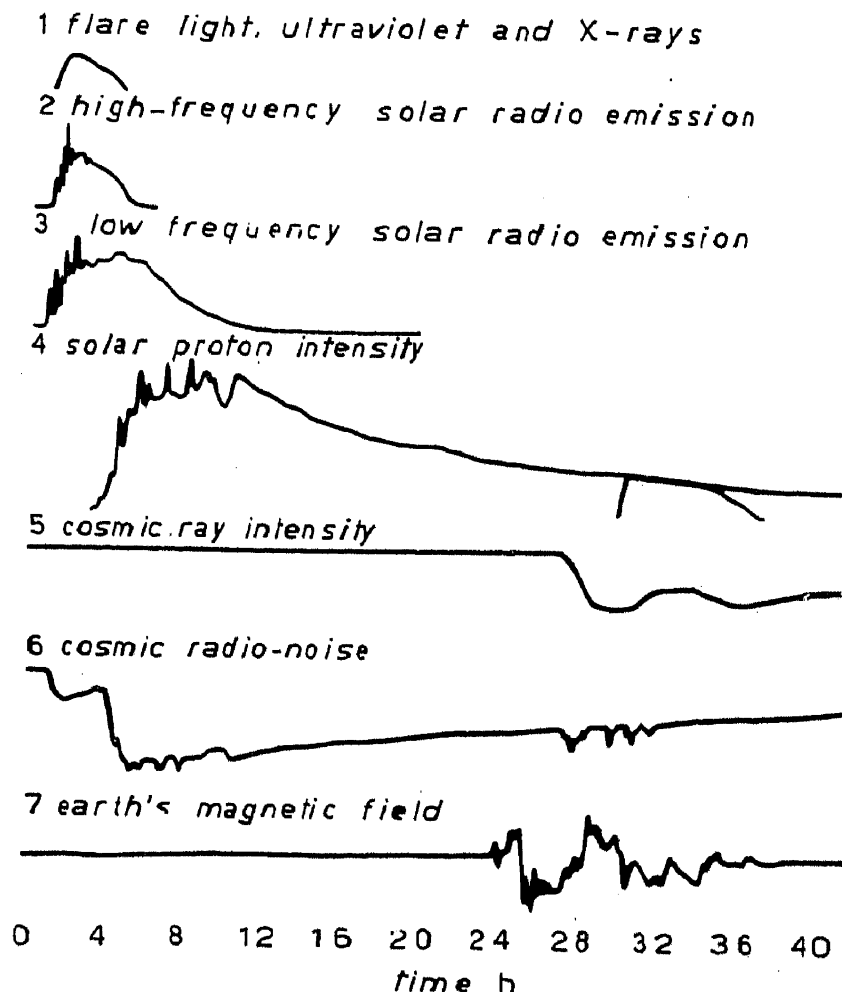


Fig. 2. - The electromagnetic radiation lasts up to four hours after the event. Ultra-violet and X-rays give an extra ionization in the ionosphere, which is manifested in a sudden decrease in the general radio noise from the galaxy (compare first part of 6 below). 1) are 2) are radio emissions accompanying the electromagnetic radiation. 3) Flare-produced protons arriving at the center many days after the onset of the flare. 4) Cosmic ray intensity is influenced by the plasma beam from the sun, when it reaches the center after about one day. 5) See under 1). When high-energy protons arrive we get ionization further down in the atmosphere. (Polar Cap Absorption). 6) Currents induced by the plasma beam around the center and in the ionosphere influence the earth's magnetic field.

We must also point out that since 1956 altogether about 40 flares were observed to produce cosmic ray particles but only a few of these flares gave significant and useful recordings. The percentage of low energy particles in a flare spectrum is much higher than in the general cosmic ray flux. With the use of neutron sensitive instruments flare particles can be studied in a better way than using meson telescopes and ionization chambers, because neutron recording apparatus responds to lower energies of the cosmic ray flux than other apparatus. Furthermore many of the flare particles are not energetic enough to produce an effect at sea level, apparatus in balloons and satellites are better at observing the flare phenomena.

Let us now list the following characteristics of solar flares as regards cosmic ray intensity. There is a rapid increase in cosmic ray intensity, with maximum occurring (10÷100) minutes after the onset of the flare, as revealed by the light- and radio emission. The recovery to normal level has a time constant of the order of a day or less.

Differences in onset time (about ten minutes) have been observed at different points on the earth.

The differential rigidity spectrum of the particles, responsible for the onset and the first part of the event, is proportional to $p^{-4} \div p^{-5}$. These particles

seem to arrive more or less directly from the sun at the earth without being scattered.

From the many models which have been proposed in order to explain the observed cosmic ray increases on the earth and their time-dependence, the following two are discussed.

One model, put forward by SIMPSON *et al.* (Fig. 3) suggests a trapping region around the sun with a radius ~ 0.3 A.U. This region contains disordered mag-

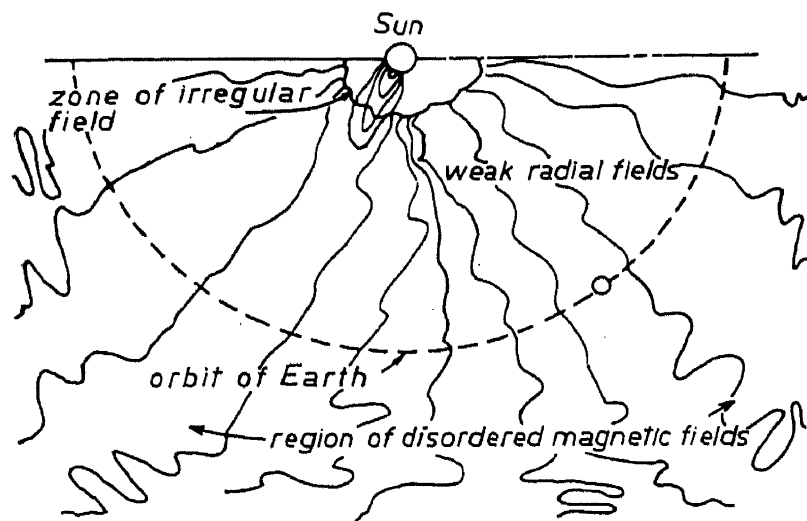


Fig. 3.

netic fields, which scatter the particles out into the solar system. Outside this region the space is essentially field-free ($B < 10^{-6}$ gauss) and particles propagate along straight lines to the earth, which means that the Firor impact zones are valid. Outside the sun-earth region there is a barrier of disordered magnetic fields, which scatter some particles back and thus have a trapping effect.

According to this picture flare particles are first scattered in all directions by the inner trapping region around the sun into field-free space. They also

leak through the barrier and propagate to the earth along straight lines and reach the Firor zones. The different onset times observed on the earth could be explained by an energy-dependent scattering near the sun.

After this first phase particles fill up the field-free region, and then are scattered by the outer barrier and approach the earth from all directions. This constitutes the indirect radiation, which decays gradually as the trapped particles diffuse through the barrier out in the galaxy.

Although this model is able to explain many of the observations, it has some disadvantages. Thus it is doubtful that the «field-free» region had a magnetic field less than 10^{-6} gauss during the event. There was for example one station, lying completely outside the Firor impact zones, which measured a very early onset time. It must in this connection be pointed out that space probes have recently measured magnetic fields in interplanetary space which are much greater than 10^{-6} gauss. Furthermore it would be difficult to explain the 27-day variation in the cosmic ray intensity if there is no real link between the sun and the earth, as a corpuscular beam with its magnetic field would constitute.

It would be interesting to know if all flares, observed by cosmic ray recording instruments can be traced to flares on the visual side of the sun. The model discussed implies that flares on the hidden side of the sun give cosmic ray increases as well.

Whereas the basic feature of the model discussed above is a field-free region between sun and earth with disordered magnetic fields outside the region, other models are characterized by an interplanetary magnetic field of a more regular structure. The following model, suggested by ALFVÉN *et al.* may in this respect represent these other theories.

During the flare of February 23, 1956 two meson telescopes in Stockholm recorded the increase. The telescopes were both tilted 30° from the vertical, one pointing in the north and the other in the south direction. This means that atmospheric effects are the same in the two telescopes. The north-pointing telescope measured during the first hour a higher increase relative to the preflare level than the south-pointing telescope did. Fig. 4.

A similar telescope arrangement in Rome indicated the same difference. The following interpretation was given.

First of all one must take into account the deflection of the particles in the earth's magnetic field. This can be done by using the results of model experiments, performed in Stockholm. In these experiments the deflection of an electron beam (imitating cosmic rays) in a magnetic dipole field (imitating the earth's field) was measured.

The results are shown on big globes, Fig. 5.

On such a globe it is possible to find the measuring direction or asymptotic direction of a given telescope so to say «outside the earth's field». This di-

rection depends on the energy of the cosmic ray particles as well as the geometrical direction of the telescope on the earth. In fact the whole solid angle of the telescope is found on the globe, *i.e.* transformed through the magnetic

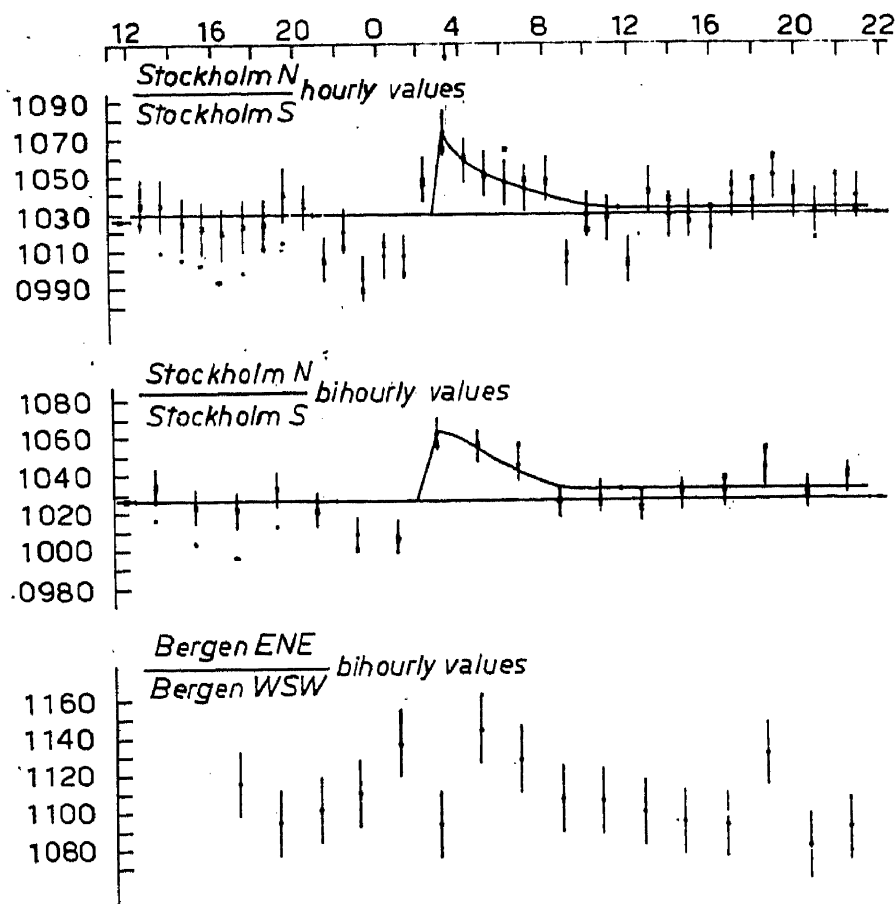


Fig. 4.

field. As regards the above mentioned telescopes in Stockholm, their transformed solid angles are seen on the globe for a particle energy of 20 GeV (Fig. 6). For lower energies the solid angles approach each other and for energies below 10 GeV they coincide more or less and are completely mixed.

From these facts and the observed difference in the two telescopes the following conclusions were drawn:

1) There was during the first part of the increase a marked anisotropy in distribution of the radiation around the earth for energies above 10 GeV. (This might also apply to energies below 10 GeV, but for these energies the measuring directions of the telescopes coincide more or less thus making it impossible in this way to measure any anisotropy in the low energy particle flux.)

2) There were consequently particle energies above 10 GeV in the flare-produced radiation. These points are supported by other measurements.

3) As the asymptotic direction of the north-pointing telescope makes an angle of $\sim 55^\circ$ with the sun-earth line particles from the sun must have been deflected through an angle of $\sim 55^\circ$, either by a magnetic field or by scat-

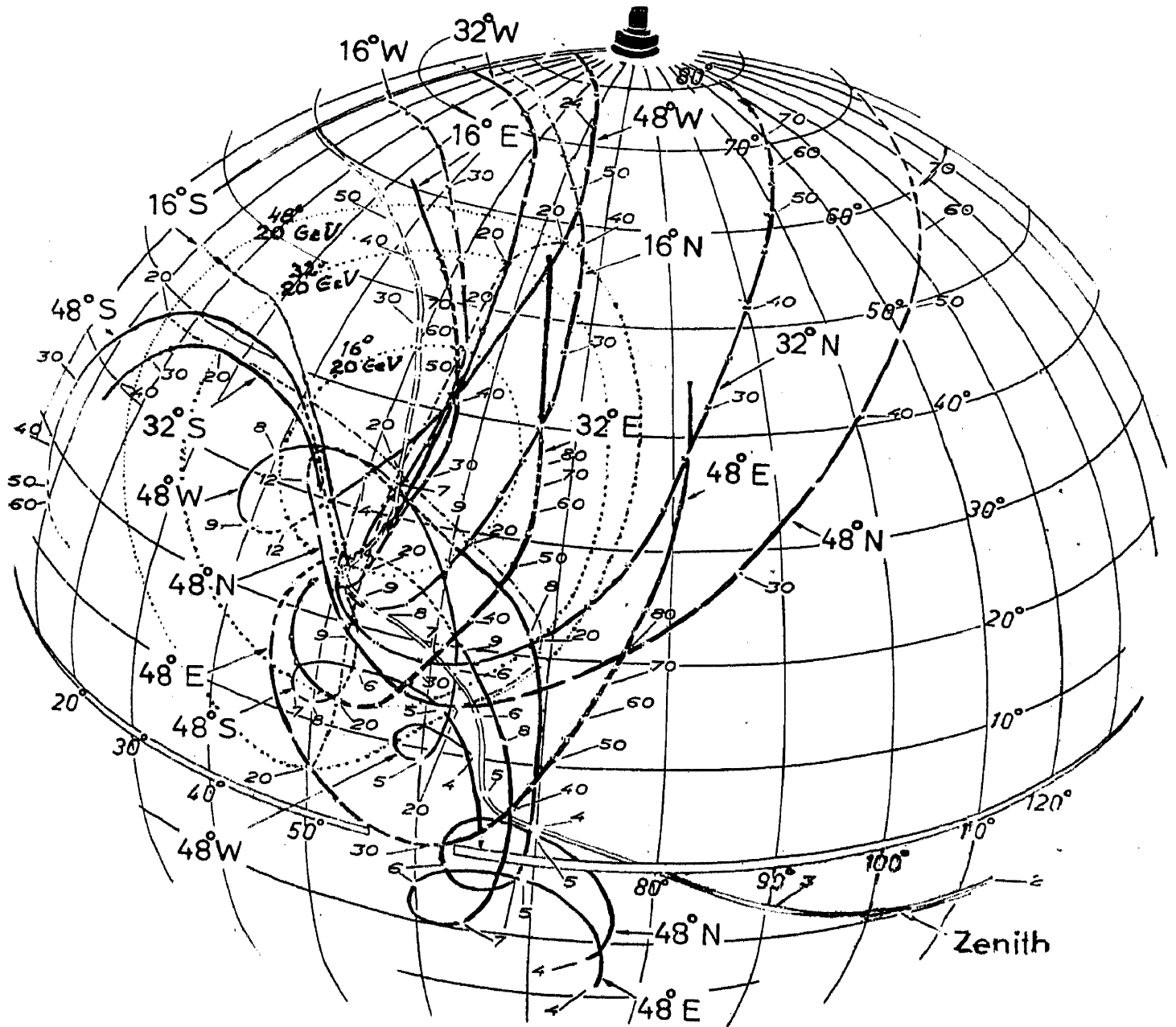


Fig. 5.

tering. Due to the big radius of curvature, scattering is not likely. Taking the energy 10 GeV this means an average field perpendicular to the ecliptic of the order of $0.5 \cdot 10^{-5}$ gauss.

In order to explain the different onset times of the increase and the high abundance of low energy particles (~ 1 GeV) in the radiation, a radial mag-

netic field (tongue-shaped) already existing before the flare and produced by an earlier event was postulated, Fig. 7.

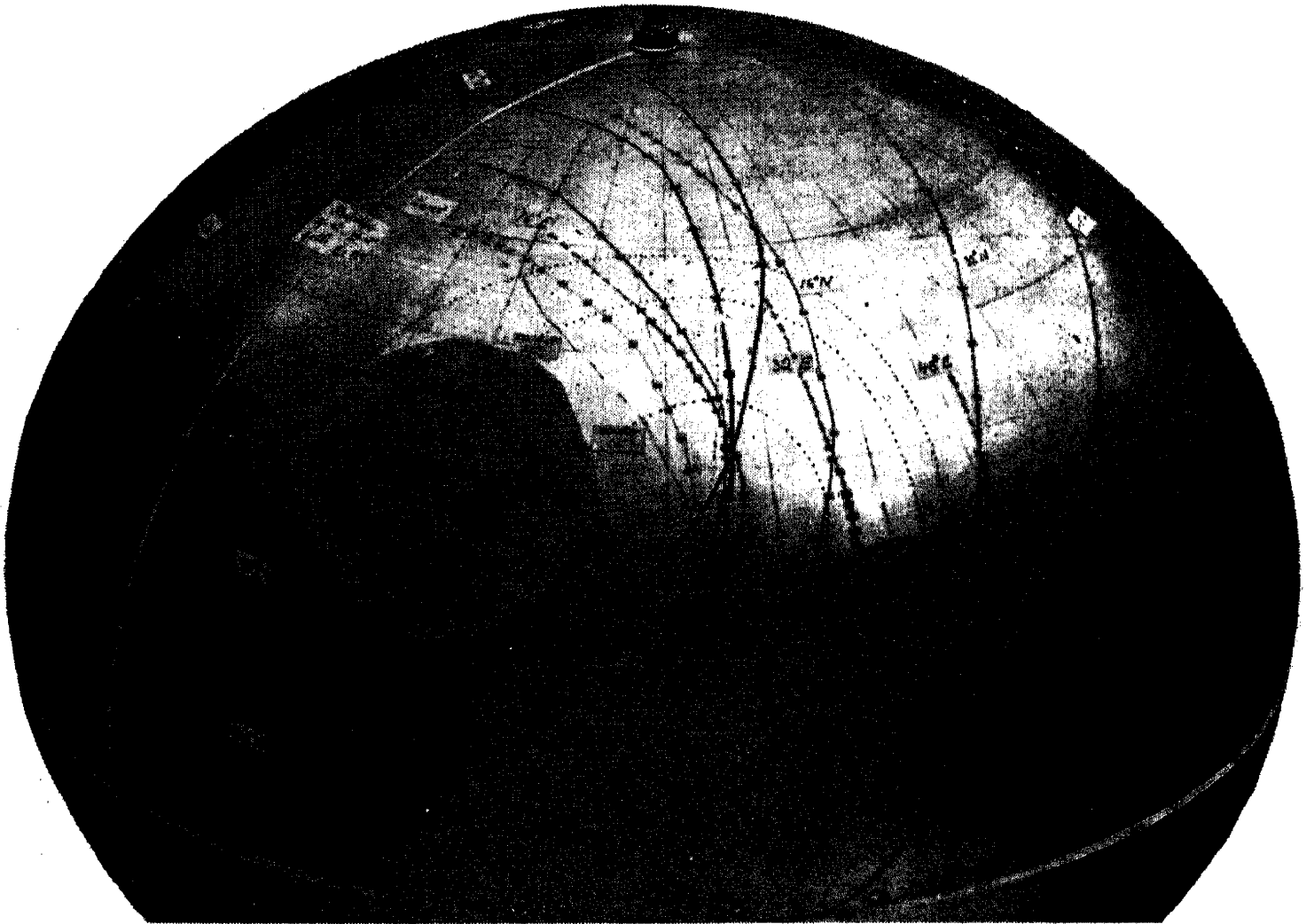


Fig. 6.

The propagation mechanism could have been as follows.

Particles of highest energies pass more or less directly from the flare to the earth and are deflected in the general solar magnetic field.

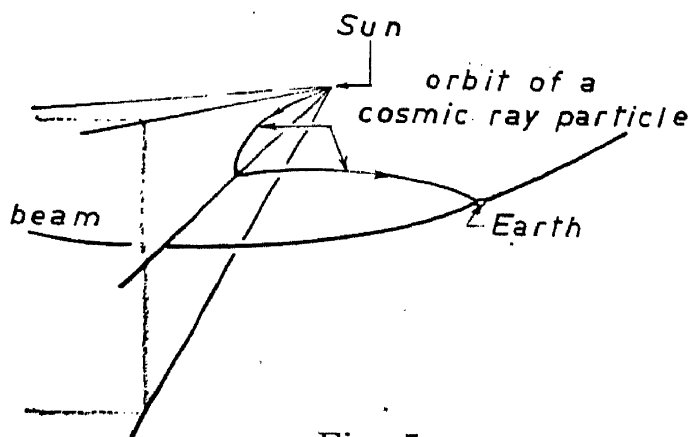


Fig. 7.

Particles of lower energies are scattered from the flare into the existing radial field. The scattering region near the sun could be similar to the one suggested by SIMPSON *et al.*

In the radial field particles move straight or spiral along the lines of force. Due to turbulent regions in the beam they are scattered out of the beam and deflected in the general solar magnetic field before reaching the earth. The

different pathlengths between sun and earth can according to Eckhart explain the different onset times, Fig. 8.

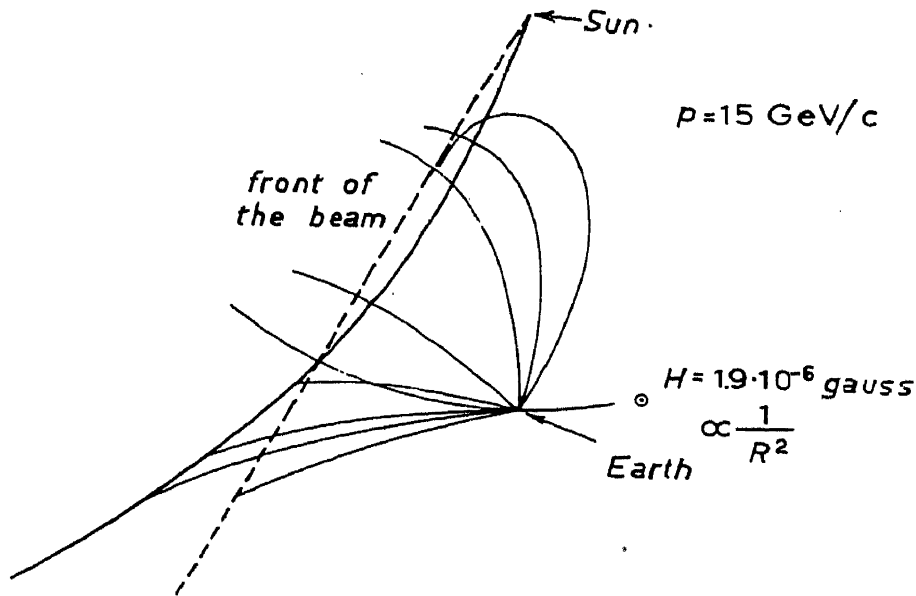


Fig. 8.

This model however has some drawbacks.

Firstly it is difficult to explain the isotropic distribution of flare particles in space as was observed during the later part of the increase. Only a limited region of the earth (facing the beam approximately) is allowed.

Secondly the magnetic field of the beam must be large with strong gradients in order to give the necessary scattering and nothing is known about this.

4. - Acceleration of particles in the solar system.

As mentioned in the beginning of the lecture there are many theories about the origin of the cosmic radiation. I understand that Prof. HOYLE will later during this course give a talk about this matter. In the following the basic ideas of the theory for a «local origin», i.e. within our solar system will be given. The theory which is worked out by ALFVÉN is, however, quite general and can be applied anywhere where we have charged particles moving in varying magnetic fields.

We consider first a charged particle moving in a magnetic field B and its motion in a plane perpendicular to B with particle momentum $p_{\perp}$.

We have:

$$r = \frac{p_{\perp}}{eB}.$$

The magnetic flux through the orbit is:

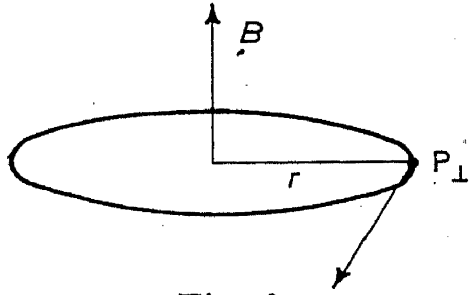


Fig. 9.

$$\Phi = \pi \cdot r^2 \cdot B.$$

Now let the magnetic field vary slowly with time:

$$\frac{d\Phi}{dt} = \pi r^2 \cdot \frac{\partial B}{\partial t} + 2\pi r \frac{\partial r}{\partial t} \cdot B.$$

The particle energy is changed due to the time derivative of Φ . Under the assumption that the energy change per turn is small the last term can be neglected and we write:

$$\frac{d\Phi}{dt} \approx \pi r^2 \cdot \frac{dB}{dt},$$

which means an electric field along the circular orbit:

$$E \approx \frac{1}{2\pi r} \cdot \pi r^2 \cdot \frac{dB}{dt} \approx \frac{r}{2} \cdot \frac{dB}{dt}$$

now

$$\frac{dp_{\perp}}{dt} = eE = e \cdot \frac{r}{2} \cdot \frac{dB}{dt}$$

and together with

$$r = \frac{p_{\perp}}{eB}$$

we get:

$$\frac{dp_{\perp}}{dt} = \frac{p_{\perp}}{2B} \cdot \frac{dB}{dt}.$$

and after integration:

$$(1) \quad \frac{p_{\perp}^2}{B} = \text{const.}$$

The particle momentum parallel to the magnetic lines of force is not influenced by the changing field:

$$(2) \quad p_{\parallel} = \text{const.}$$

We now assume that particles are moving isotropically in a homogeneous magnetic field B_0 between two perfectly reflecting planes. The magnetic field changes in the following way, Fig. 10.

At the time t_1 . We consider particles with momentum p_0 moving isotropically. That means a distribution of the momentum vectors so that as an average:

$$p_{\perp 0}^2 = \frac{2}{3} \cdot p_0^2,$$

$$p_{\parallel 0}^2 = \frac{1}{3} \cdot p_0^2,$$

The magnetic field increases during the time τ_0 until the magnetic field is B_1 . According to eqs. (1) and (2) we get: at the time t_2 :

$$\left. \begin{aligned} p_{\perp 1}^2 &= \frac{2}{3} \cdot k \cdot p_0^2 \\ p_{\parallel 1}^2 &= \frac{1}{3} p_0^2 \end{aligned} \right\} \quad p_1^2 = p_{\perp 1}^2 + p_{\parallel 1}^2 = \left(\frac{1}{3} + \frac{2}{3} k \right) p_0^2,$$

where $k = B_1/B_0$.

Now we suppose that in the magnetic field there are a number of small-scale disturbances, irregularities, which scatter particles in all directions. This means a redistribution of the energy or the momentum vector so that at the time t_3 we have:

$$\left. \begin{aligned} p_{\perp 3}^2 &= \frac{2}{3} \left(\frac{1}{3} + \frac{2}{3} k \right) p_0^2 \\ p_{\parallel 3}^2 &= \frac{1}{3} \left(\frac{1}{3} + \frac{2}{3} k \right) p_0^2 \end{aligned} \right\},$$

when this distribution is established. The magnetic field decreases back to B_0 and at the time t_4 we have:

$$\left. \begin{aligned} p_{\perp 4}^2 &= \frac{2}{3k} \left(\frac{1}{3} + \frac{2}{3} k \right) p_0^2 \\ p_{\parallel 4}^2 &= \frac{1}{3} \left(\frac{1}{3} + \frac{2}{3} k \right) p_0^2 \end{aligned} \right\} \quad p_4^2 = \left(\frac{2}{9k} + \frac{4}{9} + \frac{1}{9} + \frac{2k}{9} \right) p_0^2 = \left(\frac{5}{9} + \frac{2k}{9} + \frac{2}{9k} \right) p_0^2$$

or

$$p_4 = \alpha \cdot p_0,$$

where

$$\alpha = \sqrt{\frac{5}{9} + \frac{2k}{9} + \frac{2}{9k}}.$$

If we put $k = B_1/B_0 = 10$, $\alpha = 1.6$.

The result of the process is consequently a net gain in energy. In order to get a gain it is necessary to assume scattering, so that part of the energy, taken from the increasing magnetic field between t_1 and t_2 is stored in kinetic

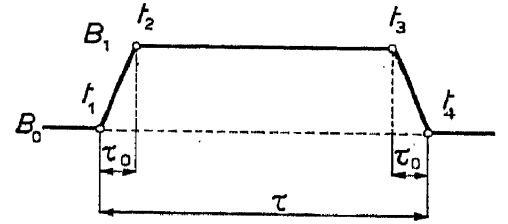


Fig. 10.

energy parallel to the magnetic field lines and thus cannot be converted back into magnetic energy when the field goes back to B_0 . If there were no scattering the particle energy would be the same at times t_1 and t_4 looking at $\alpha = \sqrt{(5/9 + (2k/9) + (2/9k))}$ it is clear that the result is the same whether the magnetic field increases from B_0 to B_1 or decreases from B_0 to B_1 . In the first case $k_1 = B_1/B_0$ and in the second case $k_2 = B_0/B_1$.

If the process is repeated the momentum will each time increase by the factor α and after n times we have by this «magnetic pumping»:

$$p_n = \alpha^n \cdot p_0.$$

5. - Discussion of the model.

The discussion above shows that changing magnetic fields transfer energy to charged particles if the scattering conditions are full filled, *i.e.* the redistribution time is much greater than τ_0 . If this is not the case the redistribution takes place more or less during the change of the magnetic field and the net gain becomes smaller. However we shall not discuss this now, but instead compare the spectrum which can be deduced from the theory with the empirical cosmic ray spectrum. Both follow a power law (p^n), the former with $n = -1$ and the latter with $n = -2.6$. A possible way of obtaining the observed value, may according to ALFVÉN, be to assume that particles are accelerated to a spectrum with $n = -1$ and then diffuse from the region of injection and acceleration so that at a certain point the density is such as to give us the observed spectrum.

In a following lecture G. G. FÄLTHAMMAR will give more details and the mathematical background of the magnetic pumping mechanism.

Acceleration of Cosmic-Ray Particles by Magnetic Pumping.

C.-G. FÄLTHAMMAR

Royal Institute of Technology - Stockholm

1. - Introduction.

The idea that the mechanism of acceleration of cosmic rays is of an electromagnetic nature was first suggested by SWANN [1] in 1933. The type of acceleration considered by SWANN is essentially the same as has later been used in the betatron high-energy accelerator and is now generally referred to as betatron acceleration. Because the high electrical conductivity of the cosmical plasma precludes electrostatic acceleration mechanisms, all current theories assume the interaction of charged particles with magnetic fields to be responsible for the acceleration. But there is no general agreement about the details of the mechanism (or mechanisms) or about where in space the acceleration takes place. In the following we shall analyse one of the acceleration mechanisms that have been proposed.

ALFVÉN [2, 3] and ALFVÉN and ÅSTRÖM [4] have proposed a mechanism based on betatron acceleration in combination with scattering of the particles against small-scale magnetic irregularities. Owing to the combination with scattering the betatron acceleration produces a systematic energy increase in fluctuating magnetic field even if the average field does not increase. The mechanism has become known as «magnetic pumping». It was originally applied to a theory of cosmic-ray acceleration in interplanetary space, but it may work also in interstellar space, if fluctuating magnetic fields are present [5]. (A similar process on laboratory scale has been considered as a possible mechanism for heating plasma to the extreme temperature aimed at in thermonuclear research [6, 7].)

The scattering that plays an important role in the acceleration has also the effect of making the particles diffuse in space during the acceleration. The combined action of these processes may explain not only the acceleration of particles to cosmic-ray energies but also the energy (momentum) spectrum that characterizes the cosmic radiation observed at the earth.

The pumping mechanism is described below, and a differential equation for the distribution of particles in space and momentum is derived. The equation is solved for particles originating at a point source in the case where the acceleration and diffusion are space-independent. From the solution are obtained the results of Alfvén's theory of the origin of cosmic rays.

2. - Dynamics of charged particles in magnetic fields.

The motion of a charged particle in a homogeneous magnetic field is, in the absence of other forces, a superposition of a circular motion in the plane perpendicular to the field lines and a uniform motion along the field lines. The guiding centre is constrained to move along the line of force where it was situated initially [8].

Consider a particle whose momentum vector $\mathbf{p}$ has the components $\mathbf{p}_\perp$ and $\mathbf{p}_\parallel$ perpendicular and parallel to the magnetic field. If the field strength changes slowly in the sense that the relative variation during one gyro-period is small, the moduli of the momentum components obey the relations

$$(1) \quad \frac{p_\perp^2}{H} = \text{const}$$

$$(2) \quad p_\parallel = \text{const.}$$

The component $p_\parallel$ is invariant and the value of $p_\perp$ at any particular instant t depends on its value at some earlier instant t_0 and the values of H at t_0 and t (but not on the values taken by H at times between t_0 and t). Thus in an increasing magnetic field the component of momentum perpendicular to the field (and consequently the total momentum) increases. This is the well-known betatron acceleration.

If, however, in the otherwise homogeneous magnetic field there are local irregularities which are not large compared to the radius of gyration, the behaviour of the particles is different. During the passage of a particle through such an irregularity, eqs. (1) and (2) are not valid. The momentum components change in an irregular way, and as a consequence of the change in $\mathbf{p}_\perp$ the guiding centre is displaced to another line of force [8]. The motion along the lines of force will keep the particle from colliding repeatedly with the same irregularity and instead make it encounter other irregularities. The collisions therefore cause uncorrelated changes in the momentum vector of each particle and tend to randomize the distribution of momentum direction of a group of particles. The displacements of the guiding centre are equally uncorrelated and make the guiding centre of each individual particle per-

form a kind of two-dimensional random walk across the lines of force. Thus for a group of many particles the effects of small-scale irregularities in the magnetic field will be:

1) A tendency towards equipartition of momentum among the three degrees of freedom. This implies that, on the average,

$$(3) \quad p_{\perp}^2 = 2p_{\parallel}^2.$$

2) A spatial diffusion of guiding centres across the lines of force (each guiding centre performing a random walk).

3. - Mechanisms for systematic momentum increase.

In a magnetic field, which fluctuates in time without increasing systematically, the betatron acceleration alone gives no net momentum increase. But if by some mechanism there takes place a redistribution of momentum between the different components, the total momentum p can under certain conditions increase systematically, even if the fluctuating magnetic field H does not. This effect is referred to as magnetic pumping. The redistribution of momentum can take place by collisions with small-scale magnetic inhomogeneities as discussed above. We shall first consider a special type of field variation in order to bring out in a simple way the essential features of the magnetic pumping. In this part we shall follow the presentation given in reference [3].

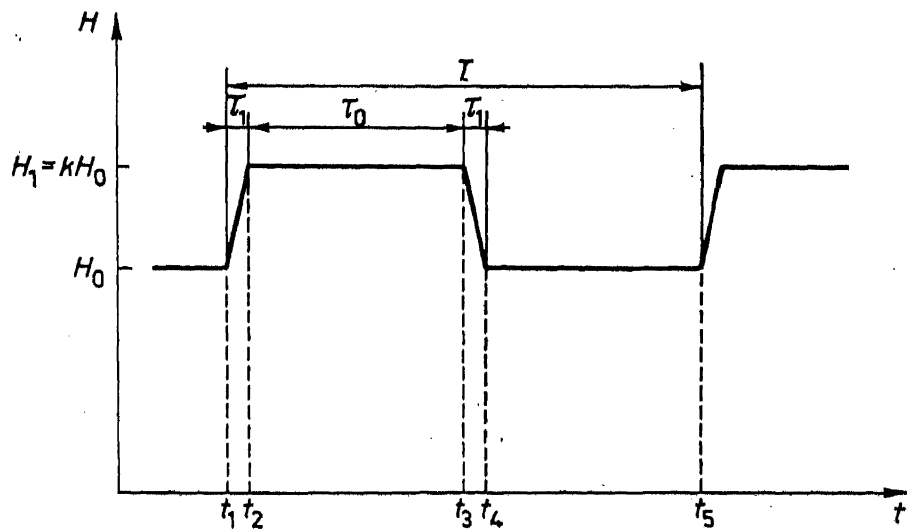


Fig. 1. - Temporal variation of the magnetic field in Alfvén's model of magnetic pumping.

Consider a magnetic field that varies in the manner shown in Fig. 1 and assume that the isotropic momentum distribution is restored at a rate cha-

racterized by the relaxation time τ_2 , which satisfies the following condition (cf. Fig. 1):

$$(4) \quad \text{gyro-period} \ll \tau_1 \ll \tau_2 \ll \tau_0.$$

Consider a set of particles which have at some instant the square momentum p_0^2 equally distributed between the three degrees of freedom. In Fig. 1 are indicated five instants of time during the cycle of field fluctuation. The processes taking place in the corresponding time intervals are the following:

In the short time interval (t_1, t_2) the perpendicular component of momentum increases in the proportion $k^{\frac{1}{2}}$, but no redistribution has time to take place. We are left at t_2 with an anisotropic momentum distribution with a total energy larger than the initial. In the next interval (t_2, t_3) isotropy is restored. This means that some of the energy gained during the betatron acceleration is poured into the parallel component of momentum, where it will stay intact during the subsequent field decrease. The field decrease in the interval (t_3, t_4) reduces again the perpendicular momentum in the proportion $k^{-\frac{1}{2}}$. But since the perpendicular momentum component is less than before the redistribution, the loss of energy during the decrease of field strength is less than the gain during the increase. Without further change in energy the distribution then becomes again isotropic during the interval (t_4, t_5) . Therefore at t_5 , after a full cycle of field fluctuation, the total momentum (p_1) is larger than that at the beginning (p_0). The corresponding quantitative consideration is straightforward and is given in Table I. There we find that the resultant momentum

TABLE I. - Average values of the momentum components and the total momentum at various instants during one cycle of the magnetic field fluctuation shown in Fig. 1.

t	$p_{\parallel}^2$	$p_{\perp}^2$	p^2
t_1	$\frac{1}{3} p_0^2$	$\frac{3}{2} p_0^2$	p_0^2
t_2	$\frac{1}{3} p_0^2$	$\frac{2k}{3} p_0^2$	$\left(\frac{1}{3} + \frac{2k}{3}\right) p_0^2$
t_3	$\frac{1}{3} \left(\frac{1}{3} + \frac{2k}{3}\right) p_0^2$	$\frac{2}{3} \left(\frac{1}{3} + \frac{3k}{3}\right) p_0^2$	$\left(\frac{1}{3} + \frac{2k}{3}\right) p_0^2$
t_4	$\frac{1}{3} \left(\frac{1}{3} + \frac{2k}{3}\right) p_0^2$	$\frac{2}{3k} \left(\frac{1}{3} + \frac{2k}{3}\right) p_0^2$	$\left(\frac{5}{9} + \frac{2k}{9} + \frac{2}{9k}\right) p_0^2$
t_5	$\frac{1}{3} \left(\frac{5}{9} + \frac{2k}{9} + \frac{2}{9k}\right) p_0^2$	$\frac{2}{3} \left(\frac{5}{9} + \frac{2k}{9} + \frac{2}{9k}\right) p_0^2$	$\left(\frac{5}{9} + \frac{2k}{9} + \frac{2}{9k}\right) p_0^2$

after one cycle is

$$(5) \quad p_1 = \frac{1}{3} \left(5 + 2k + \frac{2}{k} \right)^{\frac{1}{2}} p_0 = \beta p_0$$

The factor β is unity when k is unity, and in all other cases β is larger than unity.

After a time t , long compared to the fluctuation period T , approximately t/T cycles have passed, and the momentum has increased to

$$(6) \quad p(t) = \beta^{t/T} p_0 = e^{t/\tau} p_0,$$

where

$$(7) \quad \tau = \frac{T}{\ln \beta}.$$

Thus the momentum of the particles fluctuates, but in the mean it increases at a rate given by

$$(8) \quad \frac{dp}{p} = \frac{dt}{\tau}$$

(For individual particles the acceleration rate will differ from the mean acceleration rate given by eq. (8), and this will produce a spread in momentum of particles having initially the same momentum. The implication of this will not be considered in the present treatment.)

The special type of field variation assumed above is not essential for the mechanism to work. The problem of other types of field fluctuations, especially sinusoidal fluctuations has been studied by SCHLÜTER [6] (in connexion with a laboratory plasma, where interparticle collisions cause the momentum redistribution) and independently by MURTY and VARMA [9]. Thus, for situations where the distribution is not too far from isotropic, SCHLÜTER has shown that

$$(9) \quad \frac{d}{dt} \frac{W_{\perp}^2 W_{\parallel}}{H^2} = \frac{\gamma}{3} \frac{W_{\perp} (2 W_{\parallel} - W_{\perp})^2}{H^2},$$

where $W_{\perp}$ and $W_{\parallel}$ are the energy contributions of the perpendicular and parallel motions respectively and γ is the collision frequency. The right-hand side is always positive except in the equilibrium state, where $W_{\perp} = 2 W_{\parallel}$. Thus if we start from an equilibrium state and let the magnetic field vary in an arbitrary manner and finally return to its original value, the equilibrium which is established after this process has always a higher energy than the initial state.

In the special case of harmonic field fluctuations it has been shown [6, 9] that the momentum increases in the average exponentially so that

$$(10) \quad \frac{dp}{p} = \frac{dt}{\tau}.$$

The time constant τ depends on the relative amplitude $(\Delta H)/H$ and period T of the field fluctuation as well as on the relaxation time τ_2 for restoration of isotropy. SCHLÜTER [6] gives the approximate relation

$$(11) \quad \frac{1}{\tau} \sim \left(\frac{\Delta H}{H} \right)^2 \frac{\gamma}{1 + \gamma^2/\omega^2}.$$

where ω is the angular frequency of the fluctuation. This shows that the maximum relative increase of momentum per period of fluctuation occurs when

$$(12) \quad \gamma \approx \omega.$$

It is also evident that if the time of redistribution is too small or too large in relation to the period of fluctuation, the efficiency of the acceleration is decreased. This fact will have a selective effect resulting in preferential acceleration of such particles as have a favourable time of redistribution.

In the cosmical applications the fluctuations are of course neither of the type shown in Fig. 1, nor sinusoidal, but it should still be reasonable to expect that in the mean the momentum of the particles increases systematically according to a relation of the form

$$(13) \quad \frac{dp}{p} = \frac{dt}{\tau},$$

where τ may be a function of both position and momentum of the particle.

The analysis given above of the simple type of field variation illustrated in Fig. 1 brings out clearly a fundamental requirement for systematic energy (or momentum) increase, namely the existence of a «reservoir» in which energy is stored. In the processes discussed so far energy was stored as motion parallel to the magnetic field, but this is not the only possibility. In reference [3] is given a detailed discussion of a model in which energy is stored in an adjacent space volume, where the magnetic field does not fluctuate.

Injection. — It was mentioned above that not all charged particles are accelerated. A precise prediction of the selective action of the acceleration would require a much more detailed knowledge of the magnetic field in space than

is available at present. But, as has been suggested by ALFVÉN [3], it seems reasonable to assume that the acceleration will preferentially act on particles that have already a relatively high energy. Therefore energetic particles have to be « injected ». We shall concentrate on the case where the injection mechanism is confined to a small region in space (point-source).

4. – The distribution function.

Differential equation for the distribution function. – The combined action of acceleration by magnetic pumping and simultaneous diffusion determines what kind of spectrum will in a stationary state be observed at any given distance from the source. To find this spectrum we shall study the distribution function $f(\mathbf{r}, p)$ of the particles in the space subtended by the components of the position vector $\mathbf{r}$ of the guiding centre and by the modulus p of the momentum. Thus $f(\mathbf{r}, p) dV dp$ is the number of particles with momenta in the interval $(p, p+dp)$ that have their guiding centres in the volume element dV at the point $\mathbf{r}$. The processes affecting the distribution function are:

1) The acceleration according to eq. (13) causes particles in a volume element dV to pass the momentum value p at the rate of

$$(14) \quad f \frac{dp}{dt} dV = f \frac{p}{\tau_a} dV$$

particles per second, and so the number of particles in the phase space element $dp dV$ changes at the rate

$$(15) \quad dV \left[p \left(\frac{f}{\tau_a} \right)_p - (p + dp) \left(\frac{f}{\tau_a} \right)_{p+dp} \right] \approx - \frac{\partial}{\partial p} \left(\frac{pf}{\tau_a} \right) dp dV.$$

2) The diffusion across the lines of force implies a spatial flow of guiding centres. In the general case this flow proceeds at different rates at different momenta. For the particles in a momentum interval dp we can represent the flow vector by

$$(16) \quad \mathbf{i}(p) dp = - D(p, \mathbf{r}) \text{grad}(f dp),$$

where the gradient is taken in the plane perpendicular to the magnetic field. (When only the motion across the lines of force is of a diffusion type, the stationary state problem is essentially two-dimensional.)

In some applications it may be more adequate to assume the motion of the guide centres to have the character of a nearly isotropic diffusion in all

three directions of space [3]. Then the ordinary three-dimensional gradient operator enters into the expression

$$(17) \quad \mathbf{i} = -D(\mathbf{p}, \mathbf{r}) \operatorname{grad} f$$

for the flow vector.

3) Finally we may include an absorption effect reducing the number of particles in the phase space element at the rate

$$(18) \quad \frac{f(\mathbf{r}, \mathbf{p})}{\tau_b(\mathbf{r}, \mathbf{p})} dV d\mathbf{p}$$

The equation for the steady-state distribution function is obtained by putting equal to zero the resultant change of the number of particles in any phase space region. This implies

$$(19) \quad \int_V \left\{ -\frac{\partial}{\partial p} \left(\frac{pf}{\tau_a} \right) - \frac{f}{\tau_b} \right\} dV + \int_S D \operatorname{grad} f dS = 0,$$

where dS is the element of the surface S bounding V . Transforming the last term to a volume integral and observing that the relation is to be valid for any choice of the volume V , we obtain

$$(20) \quad \operatorname{div} (D \operatorname{grad} f) - \frac{\partial}{\partial p} \left(\frac{pf}{\tau_a} \right) - \frac{f}{\tau_b} = 0.$$

Let us assume the injection to occur at a single point source (which we locate at the origin), where particles of momentum p_0 are produced at the rate $N_0 \text{ s}^{-1}$. Then the steady-state distribution function is found by solving the equation

$$(21) \quad \operatorname{div} (D \operatorname{grad} f) - \frac{\partial}{\partial p} \left(\frac{pf}{\tau_a} \right) - \frac{f}{\tau_b} = 0$$

under the condition that

$$(22) \quad \frac{pf}{\tau_a} = N_0 \delta(\mathbf{r}), \quad \text{when } p = p_0.$$

Here $\delta(\mathbf{r})$ stands for the « Dirac-function » with the properties

$$(23) \quad \delta(\mathbf{r}) = 0 \quad \text{when } \mathbf{r} \neq 0$$

and

$$(24) \quad \int_{|\mathbf{r}| \geq 0} \delta(\mathbf{r}) dV = 1.$$

The case of space-independent parameters. - When the diffusion coefficient D , acceleration time τ_a and absorption time τ_b depend only on the momentum p and not on the position r of the guiding centre, an exact solution can be found. This is

$$(25) \quad \left\{ \begin{aligned} f &= \frac{N_0 \tau_a}{p} \cdot \frac{\exp \left[-\frac{r^2}{R_p^2} - A(p) \right]}{\pi R_p^2}, \\ R_p^2 &= 4 \int_{p_0}^p D \tau_a \frac{dp}{p}, \\ A(p) &= \int_{p_0}^p \frac{\tau_a}{\tau_b} \frac{dp}{p}, \end{aligned} \right.$$

In the case of three-dimensional diffusion mentioned above, we obtain instead

$$(26) \quad \left\{ \begin{aligned} f &= \frac{N_0 \tau_a}{p} \cdot \frac{\exp \left[-\frac{r^2}{R_p^2} - A(p) \right]}{(\pi R_p^2)^{\frac{3}{2}}}, \\ R_p^2 &= 4 \int_{p_0}^p D \tau_a \frac{dp}{p}, \\ A(p) &= \int_{p_0}^p \frac{\tau_a}{\tau_b} \frac{dp}{p}. \end{aligned} \right.$$

The results (25) and (26) are still very general because the momentum dependence of D , τ_a and τ_b has not been specified.

5. - Alfvén's model of the origin of cosmic rays.

General considerations. - Generation of cosmic rays in interplanetary space has been suggested by RICHMYER and TELLER [12] and by ALFVÉN [10, 11]. ALFVÉN [2, 3, 5] and ALFVÉN and ÅSTRÖM [4] have given a detailed theory to explain not only the production of cosmic rays but also the observed momentum spectrum. The theory is based on the processes of magnetic pumping and diffusion discussed above.

From various indirect evidence it has been concluded that an interplanetary field exists (a review of this has been given by BLOCK [13]), and recently this conclusion has been verified by direct space probe measurements [14]. The motion of interplanetary matter with such a high conductivity that the magnetic field is frozen into it [8] should provide both irregularities and temporal fluctuations in the field. Finally an injector exists, namely the sun, which is known to produce energetic particles in connection with solar activity. Thus the requirements for the processes of pumping and diffusion can be expected to be satisfied.

The spectrum. – To obtain a precise expression for the distribution function and so the spectrum, one should know the precise expressions for D , τ_a and τ_b , which would in its turn require very detailed knowledge of the interplanetary magnetic field. Such knowledge is not available at present, and so it is necessary to estimate on the basis of present knowledge at least the qualitative features of D , τ_a and τ_b .

The expressions suggested by ALFVÉN [3] are:

$$(27) \quad D = D_0 \left(\frac{p}{p_0} \right)^{1-\varphi} \quad (\varphi = \text{const})$$

$$(28) \quad \tau_a = \tau_0 \left(\frac{p}{p_0} \right)^\alpha \quad (\tau_0, \alpha = \text{const})$$

$$(29) \quad \tau_b = \text{const.}$$

For a relativistic particle the velocity is approximately the velocity of light independently of p , whereas the Larmor radius ϱ is proportional to p , and so the period of gyration τ_g is proportional to p . If τ_c is the average time required for the guiding centre to be displaced by an amount of the order ϱ , we have

$$(30) \quad D \approx \frac{\varrho^2}{\tau_c} \sim \frac{p^2}{\tau_c}.$$

Now, if ϑ is the number of gyro-periods in the time interval τ_c ,

$$(31) \quad \vartheta = \frac{\tau_c}{\tau_g}$$

the expression (27) for the diffusion coefficient implies

$$(32) \quad \vartheta \sim p^\varphi.$$

If the number of turns ϑ were equal for particles of all momenta, φ should be zero. The actual magnitude of φ might be within the limits given by

$$-1 < \varphi < 1.$$

By introducing the expressions (27)–(29) and the condition

$$p \gg p_0$$

into the general expression (25), we arrive at Alfvén's result

$$(33) \quad \left\{ \begin{array}{l} f = \text{const} \frac{\exp \left[-r^2 \frac{\alpha + 1 - \varphi}{4 D_0 \tau_0} \left(\frac{p_0}{p} \right)^{\alpha+1-\varphi} - \frac{\tau_0}{\tau_b} \left(\frac{p}{p_0} \right)^\alpha \right]}{\left(\frac{p}{p_0} \right)^{2-\varphi}} \quad \text{when } \alpha \neq 0 \\ \\ (34) \quad \left\{ \begin{array}{l} f = \text{const} \frac{\exp \left[-r^2 \frac{1 - \varphi}{4 D_0 \tau_0} \left(\frac{p_0}{p} \right)^{1-\varphi} \right]}{\left(\frac{p}{p_0} \right)^{2-\varphi+(\tau_0/\tau_b)}} \quad \text{when } \alpha = 0. \end{array} \right. \end{array} \right.$$

The integral of the distribution function over the whole space is

$$(35) \quad F = \int f dV = \left\{ \begin{array}{l} \frac{\text{const}}{p^{1-\alpha}} \exp \left[-\frac{\tau_0}{\tau_b} \left(\frac{p}{p_0} \right)^\alpha \right] \quad \text{when } \alpha \neq 0, \\ \\ (36) \quad \frac{\text{const}}{p^{1+(\tau_0/\tau_b)}} \quad \text{when } \alpha = 0. \end{array} \right.$$

The spectrum observed at any particular point is found by introducing the appropriate value of r into (33) or (34).

From the behaviour of the spectral function (33) at a given distance r from the point of injection, we can distinguish between the following three momentum regions:

1) The region of small momenta

$$(37) \quad p < p_1 = p_0 \left[\frac{\alpha + 1 - \varphi}{2} \left(\frac{r}{R} \right)^2 \right] \frac{1}{\alpha + 1 - \varphi}$$

$$(38) \quad R = \sqrt{D_0 \tau_0}.$$

In this region the particle density decreases extremely rapidly when the momentum decreases (cf. Fig. 2). The reason is that particles are accelerated

out of this momentum region in a time too short for them to diffuse the distance r from the injector. The quantity R that enters is, to the order of mag-

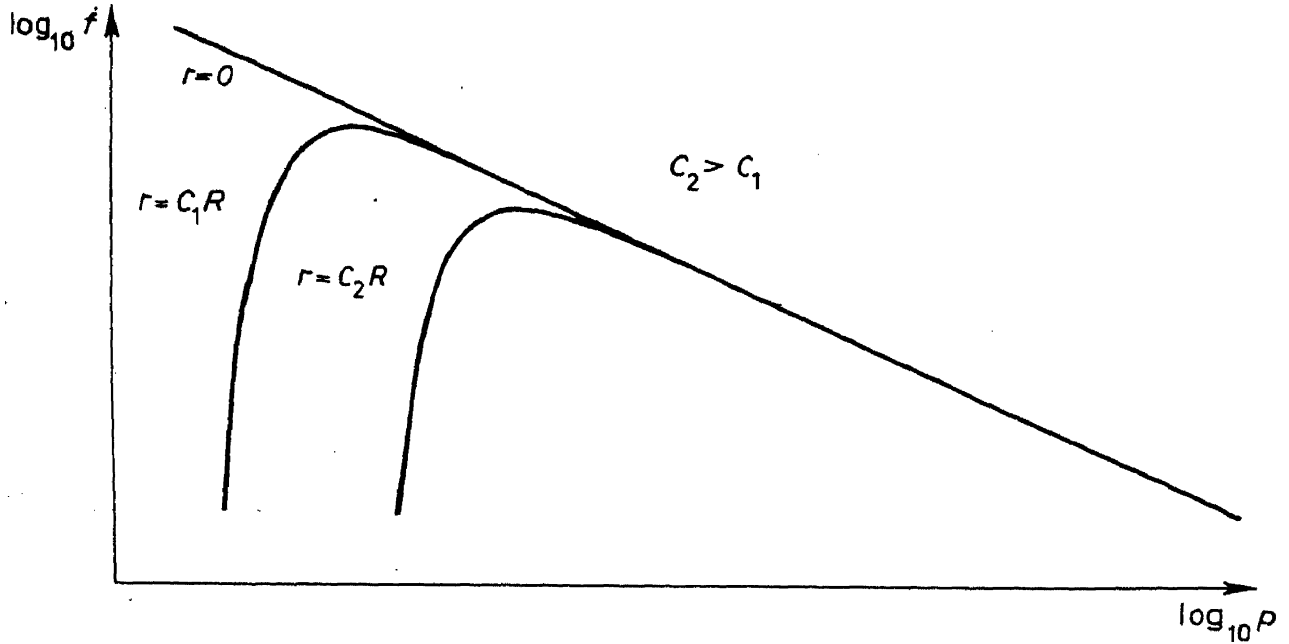


Fig. 2. - Qualitative representation of the theoretically derived momentum spectra at different distances from the particle source.

nitude, the half-width that a distribution of particles attains in the time (τ_0) required for e -folding of the initial momentum p_0 .

2) The region of large momenta

$$(39) \quad p > p_2 = p_0 \cdot \left(\frac{\tau_b}{\tau_0} \right)^{\frac{1}{\alpha}}.$$

Before the particles reach these high momentum values, they spend so long time moving through space that the probability of absorption becomes considerable. Consequently the spectrum falls rapidly (exponentially) in this region, causing an effective «cut-off» at $p \approx p_2$. (This cut-off has not been included in Fig. 2.)

3) The middle region $p_1 < p < p_2$.

In this region we have a power spectrum,

$$(40) \quad f = \frac{\text{const}}{p^{2-\varphi}}$$

(corresponding to the straight parts of the curves in Fig. 2). We also note that it is independent of the exponent α introduced in eq. (28). This independence comes about as follows: when α is large, the acceleration is slow

at high momenta and consequently high momenta are more densely populated as far as the space-integrated distribution is concerned. This is confirmed by eq. (35). On the other hand a particle requires a longer time to reach a high momentum. During this time it diffuses in space, and so fewer of the high-momentum particles remain in the neighbourhood of the injector. The two effects cancel exactly so that the value of α does not influence the spectrum.

In the interplanetary magnetic field, which must be expected to be far from homogeneous, it might be a better approximation to assume the particle motion to be of a diffusion character in all three directions of space and not only in two. The distribution function (26) would then replace (25). Introducing the eqs. (27)–(30) we obtain expressions corresponding to (31) and (32). Taking $\alpha \neq 0$ and restricting ourself to the middle momentum region where the exponentials don't play a role, we arrive once more at a power spectrum, namely

$$(41) \quad f = \frac{\text{const}}{p^{2.5+0.5\alpha-1.5\varphi}}.$$

In this case α no longer cancels exactly, but its influence is not very great.

Comparison with the observed spectrum. — In both the two-dimensional and the three-dimensional theoretical model a power law spectrum is obtained, eqs. (40) and (41). The observed primary cosmic-ray spectrum is also (above the low-energy cut-off) a power spectrum, which is given by

$$(42) \quad f = \frac{\text{const}}{p^\gamma}$$

with $\gamma = 2.5 \pm 0.2$ [15]. In the two-dimensional case, eq. (40), this would correspond to $\varphi = -0.5$. In the three-dimensional case, eq. (41), it would correspond to $\alpha = 3\varphi$, for example $\alpha = 0.3$, $\varphi = 0.1$. The implications of such values of α and φ are discussed in some detail in the paper by ALFVÉN [3] and it is concluded that they are not incompatible with observational data.

In an interplanetary field with a mean strength of 25 microgauss (which is the value found by space probe measurements [14]) the radius of curvature of a cosmic ray becomes equal to the dimensions of the planetary system when its energy is about 10^{12} eV. When the energy is larger than this, the particles travel through so large regions of space that the assumption of space-independent D and τ_a is no longer a useful approximation. Instead one should solve eq. (20) with the parameters D and τ_a having a suitable spatial variation.

If we use the common estimate of 1 microgauss as the strength of an interstellar magnetic field, we find that particles of the energy 10^{15} eV have radii

of curvature comparable to interstellar distances. At these high energies it is necessary to take also other stars into account as sources (injectors). At the highest energies the local galactic spiral arm or the whole galaxy might play the role of injector, from which cosmic rays diffuse outwards under simultaneous acceleration.

REFERENCES

- [1] W. F. G. SWANN: *Phys. Rev.*, **43**, 217 (1933).
- [2] H. ALFVÉN: *Tellus*, **6**, 232 (1954).
- [3] H. ALFVÉN: *Tellus*, **11**, 106 (1959).
- [4] H. ALFVÉN and E. ÅSTRÖM: *Nature*, **181**, 330 (1958).
- [5] H. ALFVÉN: *Tr. Moscow. Conf.* (1959), III.
- [6] A. SCHLUTER: *Zeits. Naturfor.*, **12a**, 822 (1957).
- [7] J. M. BERGER, W. A. NEWCOMB, J. M. DAWSON, E. A. FRIEMAN, R. M. KULSRUD and A. LENARD: *Phys. Fluids*, **1**, 301 (1958).
- [8] H. ALFVÉN: *Cosmical Electrodynamics* (Oxford, 1950).
- [9] G. S. MURTY and R. K. VARMA: *Phys. Rev.*, **112**, 1789 (1958).
- [10] H. ALFVÉN: *Phys. Rev.*, **75**, 11 (1949).
- [11] H. ALFVÉN: *Phys. Rev.*, **77**, 3 (1950).
- [12] R. D. RICHTMYER and E. TELLER: *Phys. Rev.*, **75**, 1729 (1949).
- [13] L. BLOCK: *Ark. för Fys.*, **14**, 179 (1958).
- [14] P. Y. COLEMAN, L. DAVIS and C. P. SONETT: *Phys. Rev. Lett.*, **5**, 43 (1960).
- [15] V. L. GINZBURG and S. I. SYROVAT'SKII: *Usp. Fiz. Nauk*, **71**, 411 (1960), English Translation: *Sov. Phys. JETP*, **3**, 504 (1961).

The Earth's Radiation Belts and the Aurora.

E.-Å. BRUNBERG

The Royal Institute of Technology - Stockholm

1. - Introduction.

One of the most interesting geophysical discoveries during the last years is that of the dense radiation belts around the earth. The discovery was made by Geiger counters in artificial satellites. Sputnik II (November 1957) showed that the counting rate increased above 500 km and at 700 km the rate was 50% above the intensity expected. In the U.S.-satellite Explorer I an increase in counting rate more than 100 times the expected value was measured and the counting rate actually blocked the counters. Professor VAN ALLEN in 1958 interpreted the high counting rate as being due to trapped particles in the earth's magnetic field. Continued satellite registrations have shown that there are two radiation belts. Their principal form is given by the first figure.

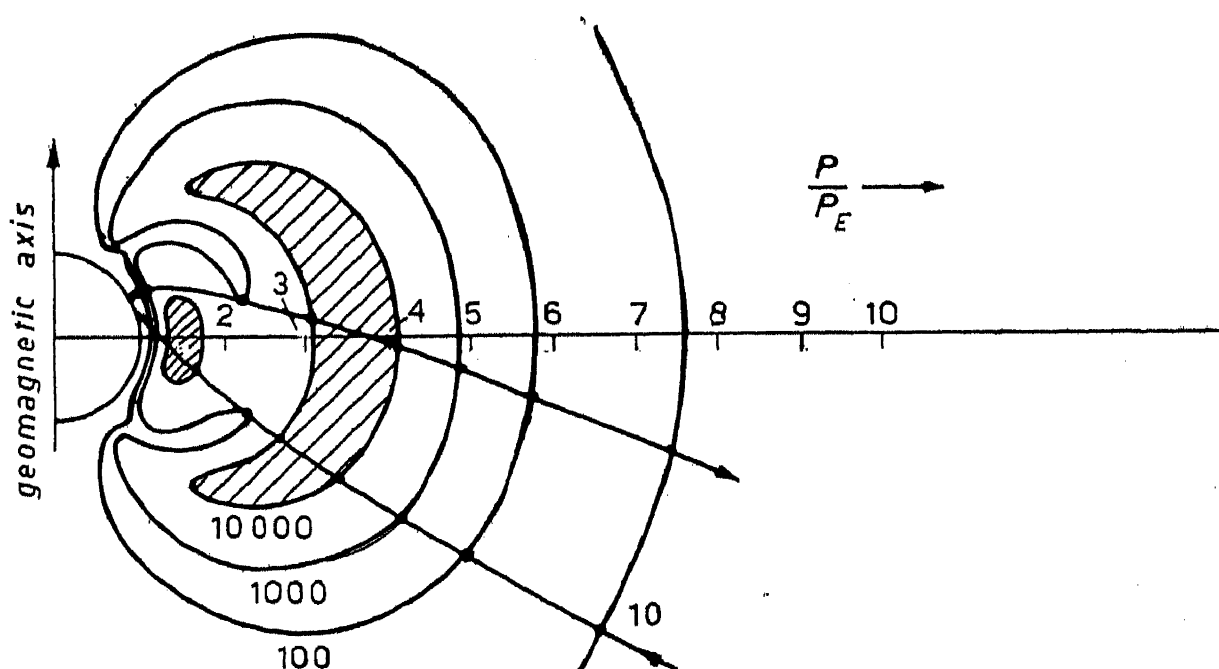


Fig. 1.

Before we go further on we shall discuss the physical and mathematical conditions for particles being trapped in a magnetic field.

In Fig. 2 we have:

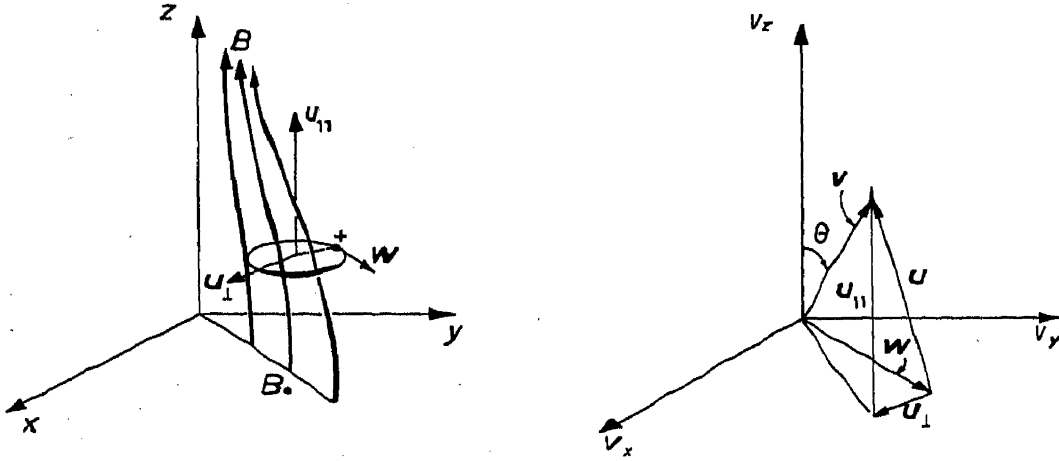


Fig. 2.

$$v = w + u \quad \text{and} \quad u = u_{\parallel} + u_{\perp},$$

where: v = total particle velocity,

w = particle velocity in the circular orbit with center C . (Guiding center),

u = velocity of the guiding center,

$u_{\parallel}$ = guiding center velocity parallel to the magnetic field lines,

$u_{\perp}$ = guiding center velocity in a plane perpendicular to the magnetic field lines.

If we assume that $u_{\perp} \ll w$, i.e. the velocity of the guiding center, C , parallel to the xy -plane is small compared to the velocity w of the particle in its circular orbit, we have (Fig. 3):

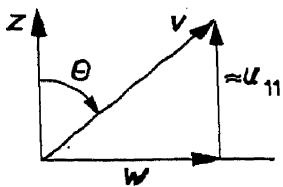


Fig. 3.

$$u_{\parallel} = v \cos \theta, \quad w = v \sin \theta.$$

The condition for $u_{\perp} \ll w$ is that $\text{grad } B = |\partial B / \partial x, \partial B / \partial y|$, is small over a distance equal to the gyro-radius. The particle moves in a spiral orbit starting with pitch angle θ . The magnetic moment of a gyrating charged particle in a magnetic field is:

$$\mu = \frac{W_{\perp}}{B},$$

where: $W_{\perp}$ = particle energy perpendicular to the magnetic field,

B = magnetic field strength.

It can be shown that the magnetic moment is approximately constant during the motion if the radius of curvature in the magnetic field is very small compared to the distances over which the magnetic field varies.

We write:

$$\mu = \frac{(mv^2/2) \sin^2 \theta}{B} \approx \text{const.}$$

Due to the velocity component $u_{\parallel}$ the guiding center of the particle moves in the z -direction into a stronger magnetic field. The condition $\mu = \text{const}$ gives the following relation between pitch angles at two different positions in the field:

$$\frac{\sin^2 \theta_0}{B_0} = \frac{\sin^2 \theta}{B}.$$

When the particle reaches a region at which $\theta = \pi/2$ it cannot proceed further and is reflected back towards the starting point. If the magnetic field is a dipole field or has a similar shape, particles starting for instance in the equatorial region will oscillate between the so called mirror points, where $\theta = \pi/2$.

Superimposed on this motion there is a drift perpendicular to the magnetic field lines, which depends on $\text{grad } B$ in planes perpendicular to B .

This drift velocity $u_{\perp}$ of the guiding center is given by:

$$u_{\perp} = \frac{W_{\perp}}{eB^2} \cdot |\text{grad } B|_{\perp}.$$

The drift velocity is perpendicular both to the magnetic field lines and to the gradient of the field.

In the equatorial plane of a dipole the drift motion is as in Fig. 4.

The magnetic field at the « mirroir point » is given by

$$B = \frac{B_0}{\sin^2 \theta_0}.$$

stronger field
 $\odot B$

$\odot B$

We apply this formula to charged particles moving in the earth's magnetic field and consider particles starting at about



weaker field



Fig. 4.

6 earth-radii in the equatorial plane. The reason for choosing this radial distance is that magnetic field lines, which start in the auroral zones, reach the equatorial plane at approximately 6 earth radii and thus particles at this distance probably play an important rôle in the auroral mechanism.

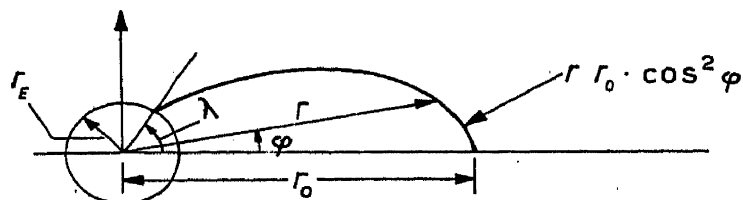


Fig. 5.

The following table gives some characteristics of trapped orbits in the earth's dipole field.

r_E = radius of the earth,

r_0 = radial distance of a magnetic field line at the geometric equator,

λ = geomagnetic latitude at which a magnetic field line starting at the equator reaches the earth,

θ_0 = pitch angle of a particle at the equator.

TABLE I.

r_0/r_E	$90 - \lambda$	λ	θ_0
1.05	77.4°	12.6°	64.0°
1.10	72.4°	17.6°	55.0°
1.50	54.6°	35.4°	27.2°
3.0	35.2°	54.8°	12.3°
6.0	24.1°	65.9°	2.8°

Considering the state of affairs in velocity space this means that particles starting from the equatorial plane with velocity vectors pointing within a cone with the full apex angle $2\theta_0$ hit the earth, or more correctly are very likely to be absorbed in the earth's atmosphere. Thus there is a «sink» in velocity space for particles starting from the equatorial plane. The solid angle of this «sink» is determined by θ_0 (Fig. 6).

We thus realize that particles, which once have been trapped in the earth's magnetic field will not stay there for ever. They are sooner or later scattered into the «sink». The life-time of the trapped radiation is determined by the scattering mechanism and the atmospheric density. The scattering may be caused by geomagnetic irregularities in the field and

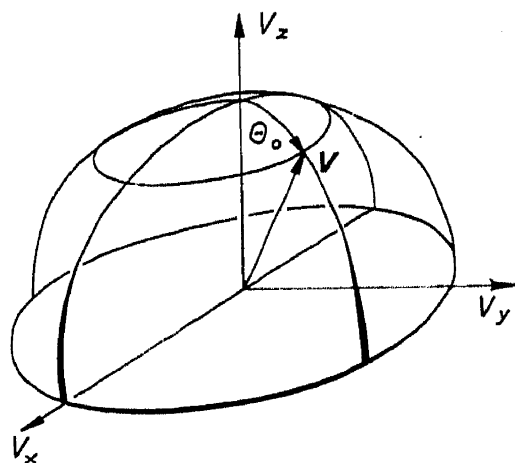


Fig. 6.

plasma oscillations as well as Coulomb scattering. Although particles are most likely to be absorbed near the mirror points, where the air density is highest, it is not certain that the scattering region here is the most important one along the field line, as has been pointed out by SINGER *et al.*

The scattering mechanism could also cause a drift of particles perpendicular to the magnetic field lines, *i.e.* a diffusion outwards from the earth as well as inwards which also influences the life time of the trapped radiation. Particles which diffuse into a stronger field are more and more likely to be scattered into a «sink» as the solid angle of the sink increases with decreasing distance from the earths.

2. - Composition of the radiation.

The following values have been given by VAN ALLEN *et al.*:

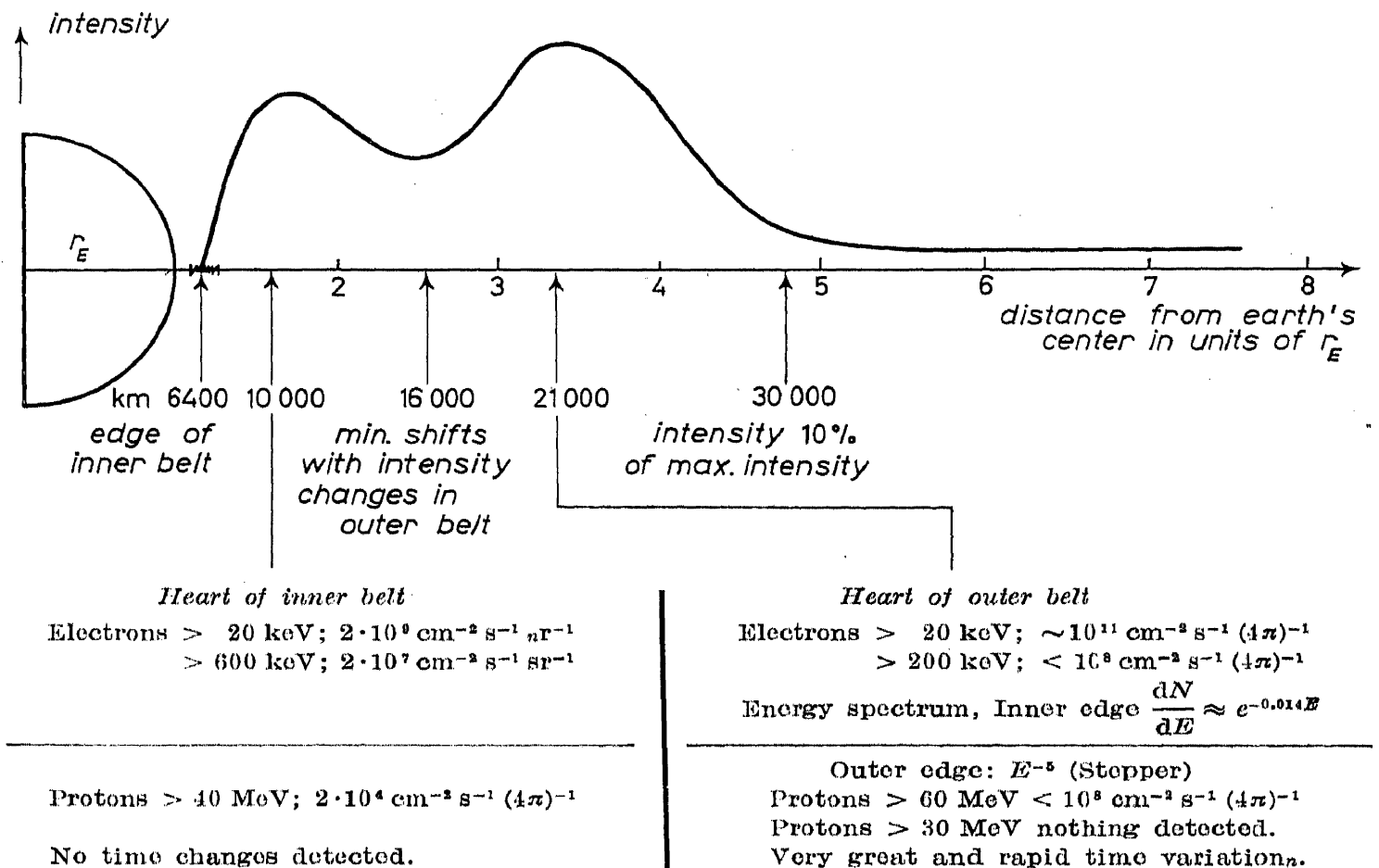


Fig. 7.

3. - Theories for the origin of the radiation.

3.1. - *Albedo particles.* SINGER has suggested that fast cosmic ray albedo neutrons could be a likely source. This suggestion depends upon the fact that neutrons, produced in cosmic ray-induced nuclear disintegrations in

the atmosphere, will move outward through the geomagnetic field without deflection and eventually undergo radioactive decay. The decay products of a neutron (half life ~ 12 minutes) are an electron, a proton and a neutrino. The kinetic energy of the proton is about the same as that of its parent neutron and the electron has a well-known β -decay spectrum with an upper limit of 782 keV for a neutron at rest. VERNOV *et al.* have independently suggested the same mechanism. They have also pointed out that the magnetic scattering would provide a limit on the life time.

It has also been suggested that the decay of μ -mesons coming from the earth's atmosphere, as a part of the cosmic ray albedo, could be responsible for the electron component in the trapped radiation, or at least for some part of the electron component.

3'2. - Injection of solar particles It has been observed that the maximum intensity in the outer zone as well as its extension varies with time. This suggests a solar origin of the outer zone. Plasma clouds or beams from the sun have been suggested as a possible source. These plasma beams may be able to penetrate sufficiently far enough into the earth's field and scatter a large number of particles of low energies into trapped orbits, primarily in the outer belt. We will come back to this later.

The inner belt seems to be much more stable than the outer belt as regards intensity and position.

We need a lot more information about the composition of the radiation and especially the irregularities of the magnetic field as well as its time dependence, before we can say anything more definite about the type of mechanism, which is responsible for the belts.

3. - The Argus experiment.

In the autumn of 1958 a geophysical experiment on global scale was conducted. Three small atomic bombs were detonated beyond the atmosphere at a location in the South Atlantic. The purpose of the experiment was to study the trapping of the electrons, produced by the β -decay fission fragments in the geomagnetic field. The experiment was suggested by CHRISTOFILOS and was called Argus. The experiment was very successful. Released electrons were trapped in the field and oscillated along the magnetic field lines between mirror points. At the same time the electrons drifted eastwards (compare Fig. 4) and created a thin electron shell around the earth. The life-time and location of this electron shell were measured by satellite- and rocket-borne instruments. Auroral luminiscence was observed at those points, which on different positions on the earth were connected with a corresponding magnetic field line. The electron shell was remarkably stable

during its life-time and no motion of the shell or change in its thickness could be detected.

Since the density of the atmosphere decreases exponentially with altitude it is evident that most of the losses should occur near the mirror points, although the scattering occurs along the whole field line and may be most important near the geomagnetic plane. The mechanism of the outer radiation belt with particles leaking into the earth's atmosphere suggests that series of geophysical effects could be explained by the existence of the belts, for instance the aurora, airglow, atmospheric heating and geomagnetic storms.

Before we go further let me first point out some observational facts about the aurora and review some of the most well-known theories.

Geomagnetic storms and aurora are very closely related to each other. Any theory which explains the aurora should also be able to explain the geomagnetic storms.

The main features of the aurora are:

1) Auroral zones. Aurora occurs in the northern and southern belts of auroral zones which are situated about $20^{\circ} \div 25^{\circ}$ from the geomagnetic poles. The auroral zones move towards lower latitudes during strong disturbances. There is some evidence for a diurnal shift in the latitude of the zone, Carl. keim-Gyllenskiöld 1886.

2) The lower limit at which aurora occurs is in general around $(80 \div 100)$ km height.

It is now generally accepted that the auroral light is caused by protons as well as electrons and that these particles give up their energy by excitation, dissociation or ionization of the air atoms and molecules. Spectroscopical measurements of the auroral light show that some percent of the light is produced by protons.

3) Relationship with solar events. There is good correlation between solar activity and strong auroral displays. If there is a solar flare then one can expect aurora about one day later on the earth, although the existence of flare is not a guarantee for an auroral display. Evidently «something» is shot out from the sun which takes about one day to reach the earth and which thus has a mean velocity of 10^8 cm/s. It is not always certain that this «something» hits the earth or comes near to it.

The first effort on a large scale to give a scientific explanation of magnetic storms and aurora was made by Birkeland (around 1908). His excellent terrella experiments gave a great amount of valuable experience. They also inspired Størmer to his great pioneering work on charged particle orbits in magnetic dipole fields, the result which is now of great importance, not only in auroral theories but even more in cosmic-ray work.

4. - The Størmer theory.

The basis for this theory is the Birkeland terrella experiment. In a big glass chamber, which was evacuated as well as possible at that time, a magnetic steel sphere (terrella) was placed. Between a cathode and the sphere a discharge was started which produced a luminous, ring-shaped pattern around the magnetic poles of the terrella. Birkeland and Størmer interpreted the experiment in the following way. The cathode emits electrons, which under the influence of the magnetic field move in complicated orbits. STØRMER was later able to show that electrons, which start from a source near the equatorial plane of the terrella, should hit its surface in circular regions around the poles. These results were applied to nature and STØRMER assumed that the sun emitted beams of electrons, which hit the earth at the auroral zones and caused the auroral light in the upper atmosphere.

There are, however, several objections to this theory.

1) Electrons penetrating not deeper into the atmosphere than to a height of $(80 \div 100)$ km, under which aurora is not observed, have energies only of the order of $(10^4 \div 10^5)$ eV, whereas orbit calculations show that an energy of 10^8 eV is needed for particles, which reach the auroral zones at about 22° from the poles. The actual electrons of energies $(10^4 \div 10^5)$ eV cannot penetrate the earth's field and reach the earth more than one or two degrees from the poles. This led STØRMER to postulate the existence of a ring current around the earth which could modify the particle motion.

2) It has been pointed out that a stream of charged particles of only one sign, strong enough to explain the terrestrial disturbances, would have an enormous space-charge and thus cannot exist.

3) The interplanetary magnetic field will deflect the electron stream.

5. - The Chapman-Ferraro theory.

This theory does not primarily explain the aurora but is of importance for later auroral theories.

The basic idea of the theory is that a stream of particles is ejected from the sun and that there is an equal number of electrons and singly charged positive ions in the beam, which thus is neutral. All particles are moving with the same velocity towards the earth. There are no electric or magnetic fields in space before the beam approaches the earth.

When the stream advances towards the earth it is influenced by the terrestrial magnetic field. If we consider the front of the stream to be plane and

infinite, the current system induced in this front by the motion into the field can be calculated. These currents modify the magnetic field, and through interactions between field and currents the stream is distorted. The modification of the earth's magnetic field, due to currents in the front of the stream are suggested by CHAPMAN and FERRARO to explain the initial phase of a magnetic storm. The theory has later been extended by MARTYN to include also the aurora which always accompany magnetic storms.

6. - The Alfvén theory.

The basic idea of this theory is that the solar beam of ionized particles carries a magnetic field with it. Due to the high degree of ionization in the beam and consequently the good conductivity, a magnetic field which is enveloped in the plasma stream cannot escape, it is a « frozen in » field.

Let us see what happens in general when a plasma stream moves in space. In the first case the beam moves with the velocity u in a magnetic field B .

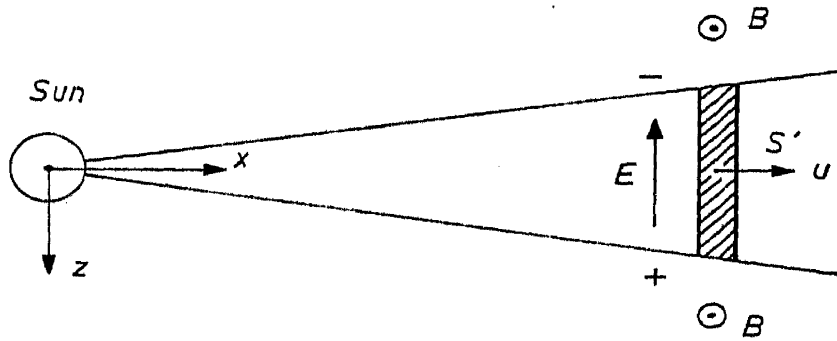


Fig. 8a.

In the sun-earth system S :

Magnetic field: B_y

Electric field: E_z

In the moving system S' :

$$B'_y = \frac{1}{\sqrt{1-\beta^2}} \left[B_y - \frac{u \cdot E_z}{c^2} \right]$$

$$E'_z = \frac{1}{\sqrt{1-\beta^2}} [E_z + u \cdot B_y],$$

$$\text{as } u \ll c \quad \frac{1}{\sqrt{1-\beta^2}} \approx 1;$$

$$\text{thus } B_y' \approx B_y,$$

$$E'_z \approx E_z + u \cdot B_y.$$

The polarization field E'_z makes positive charges accumulate on one side of the moving conductor and negative charges on the other side until the electric field from these charges neutralizes E'_z and we have: $E_z \approx -u \cdot B_y$.

Thus seen from the sun earth-system S the electric field is:

$$E_z \approx -uB_y$$

or, as can be shown, more generally:

$$\mathbf{E} = -\mathbf{u} \times \mathbf{B}.$$

In the second case we have a magnetic field B' « frozen into » the slab (S'), which moves with the velocity $\mathbf{u}$.

Relativistic transformation to the earth-sun system S gives:

$$\mathbf{E} = \frac{1}{\sqrt{1-\beta^2}} (\mathbf{E}' - \mathbf{u} \times \mathbf{B}') \approx \mathbf{E}' - \mathbf{u} \times \mathbf{B}'.$$

As $\mathbf{E}' = 0$, the electric field measured by an observer at rest in the system S is

$$\mathbf{E} \approx -\mathbf{u} \times \mathbf{B}' \approx -\mathbf{u} \times \mathbf{B},$$

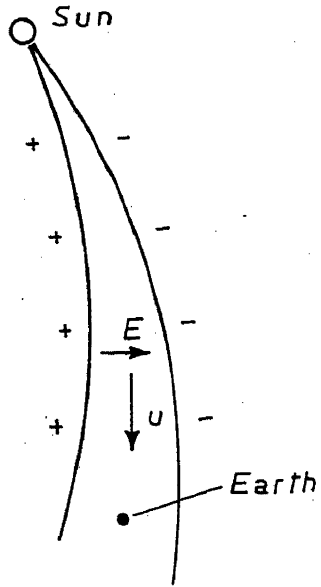


Fig. 8b.

which is the same as in the first case although now the magnetic field is moving relative to the observer.

Fig. 8b shows the situation:

The particles in the beam have low energies and move in trocoidal orbits. The electric field E is of the order of 10^{-3} V/m, which implies a magnetic field in the beam of the order of:

$$B = \frac{E}{u} \approx \frac{10^{-3}}{10^6} \cdot 10^4 = 10^{-5} \text{ gauss}.$$

It can be shown that the motion of the electrons of the beam in the combined electric field (E) and magnetic dipole field of the earth is as in Fig. 9.

There is a forbidden region at about $(5 \div 6)$ earth radii from the earth. The positive ions, however, which are supposed to have a lower temperature than the electrons, are not deflected in the same way and try to pass into the forbidden region. Just inside the border facing the sun we thus get a positive space charge. On the night-side border the electrons are uncompensated by the positive ions which do not reach this region and a negative space charge is set up. If one assumes that the only way these two regions of space charge can neutralize each other is by a discharge along the magnetic field lines of force over the polar cap, this can account for the polar disturbance and the aurora (Fig. 10). The physics and the mathematical treatment of the theory

has been investigated with model experiments by MALMFORS and BLOCK. These investigations support the theory. However, the whole theory is not proved until a convincing comparison between the different auroral phe-

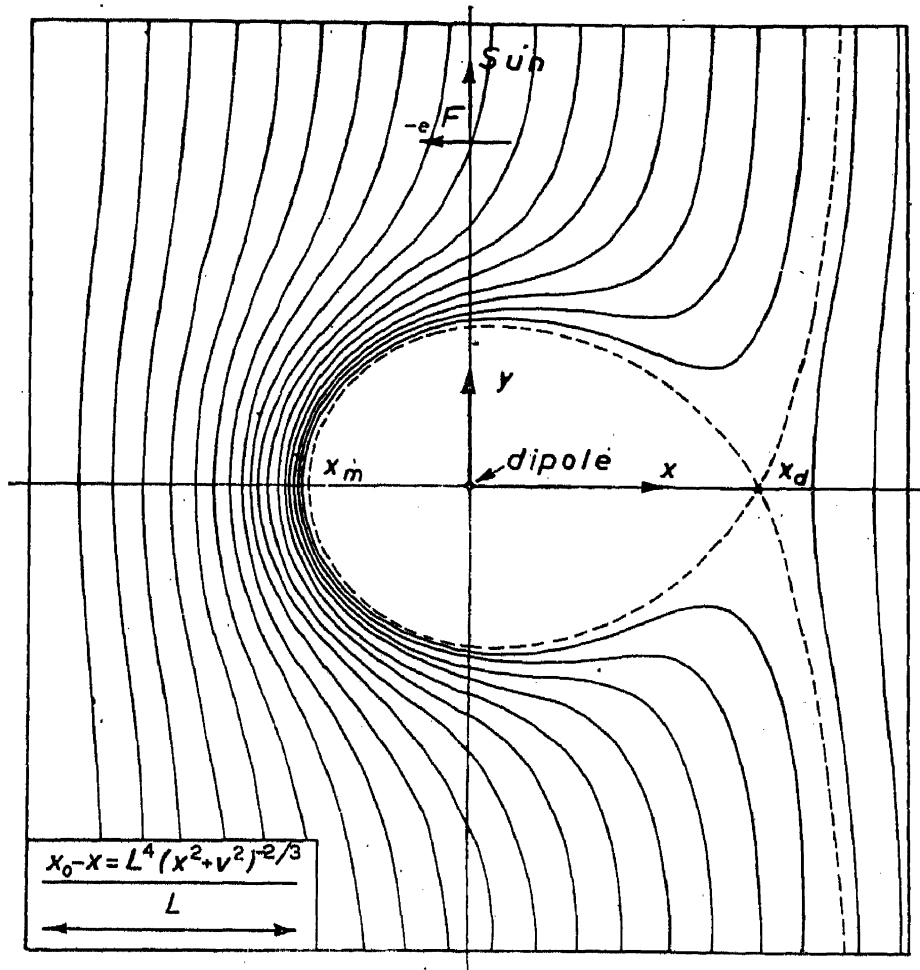


Fig. 9.

nomena in nature and in the terrella experiments have been made. This work is in progress using the results of auroral observations obtained during the geophysical year.

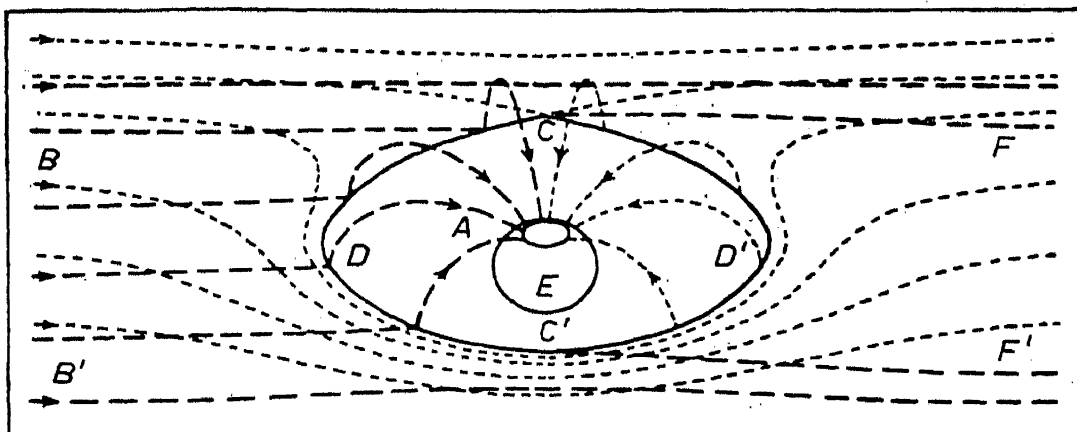


Fig. 10.

Let us now return to the radiation belts and the possibility of using them in explaining the aurora.

It has been suggested that the magnetic dipole field which constitutes a magnetic bottle for the electrons in the outer radiation belt is «squeezed» by plasma streams coming in a beam or a plasma cloud from the sun, or more correctly, electrons are scattered by magnetic disturbances and are lost in the atmosphere producing auroral light.

By considering the luminosity of aurorae the electron flux down into the atmosphere has been estimated to between $(10^9 \div 10^{11})$ electrons/cm² s where the last value corresponds to a strong aurora.

Using van Allen's data the electron density of the outer zone is about 10 electrons/cm³ and altogether there are $\sim 10^{29}$ electrons in the outer belt.

As the auroral zone is ≈ 1000 km wide its surface is:

$$\pi \cdot 6.4 \cdot 10^8 \cdot \sin 23^\circ \cdot 10^8 = 10^{17} \text{ cm}^2.$$

With an electron flux of 10^{11} electrons/cm² s this means a total electron flux of 10^{28} electrons/s and consequently the magnetic bottle is emptied in a time of the order of 10 s.

This means that particles must be fed into the outer belt continuously during the auroral display. But as it is most probable that particles in a beam generally have low energies corresponding to the temperature of the outer parts of the sun, from where they are shot out, these particles must be accelerated up to 10^4 eV in order to account for the aurora. Such a gain in energy can be understood from the Alfvén electric field model for the aurora.

Space research with rockets and satellites has now opened new possibilities to measure such things as magnetic field strength, particle density and energy distribution of particles far away from the earth.

One important investigation would be to measure the change of the earth's magnetic field in the ionosphere and along the magnetic lines of force where currents flow during aurora according to the Alfvén theory. Another interesting task is to take pictures from a satellite of the whole auroral zones during auroral displays, thus making it possible to see their whole extension and if there is a daily variation in position as predicted by the theory.

It is, indeed, a very fascinating time for those working in geophysics and cosmical physics. During decades theories have been put forward, based on measurements carried out on the earth and thus giving very uncertain, indirect information about conditions in space. For the same reason it has been difficult to reject a theory with 100% certainty and the physicists have died happy and convinced that their theories are right. This state of affairs is soon over—as regards many problems in our neighbourhood in space—but there are without doubt problems enough beyond our solar system and our galaxy that scientists might argue about until the end of mankind.

Mechanisms of V.L.F. emission.

R. E. GENDRIN

Service d'Aéronomie de l'Observatoire de Meudon, Meudon (S. et O.)

1. - Introduction.

Since the whistlers were discovered and studied on a large extent, radiation of very-low-frequency electromagnetic noise has been also recorded. This noise cannot be produced by the same mechanism (strokes) which generates the whistlers. A large and difficult, experimental and theoretical work has been achieved on this subject. We will briefly report on this work, paying more attention to the Čerenkov effect produced by extra-terrestrial particles in the exosphere, in this frequency range (0.5 to 20 kHz).

2. - Previous work.

2.1. *Experimental results.* - Noise recorded in this frequency range may be classified in two species.

First, the broad-band noise whose mean frequency and band-width are

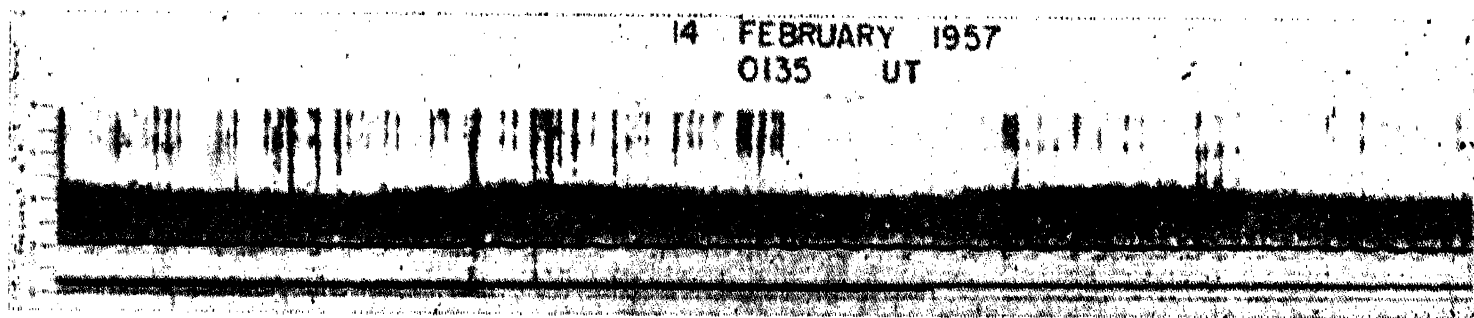


Fig. 1. - Hiss with a high-frequency cut-off near 5 kHz. After GALLET [7].

more or less constant in time for a given place [1-4]; one of these frequencies is nearly 750 Hz, and has been attributed to the gyrofrequency emission of protons [5, 6] (Fig. 1 and 2).

Second, emissions reduced to a single frequency, which may be or not constant in time, are sometimes observed. (We will call them discrete emis-

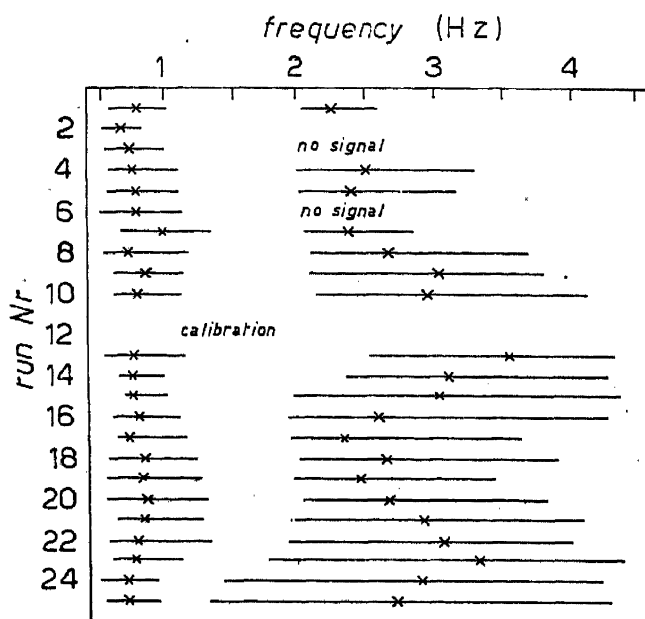


Fig. 2. — Emission around the proton gyrofrequency. Length of the lines represents the bandwidth of the recorded emission. Crosses indicate frequency of maximum intensity. Each run lasts around 20 min. After AARONS *et al.* [5].

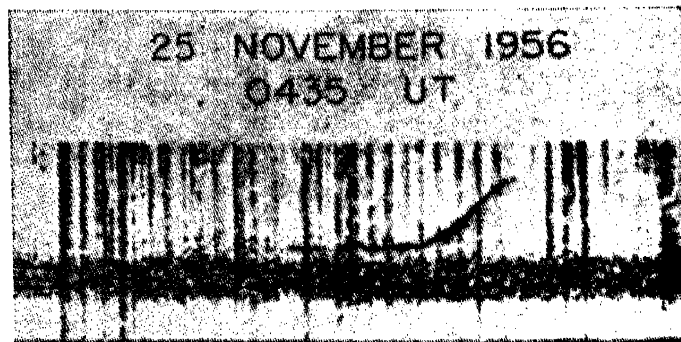
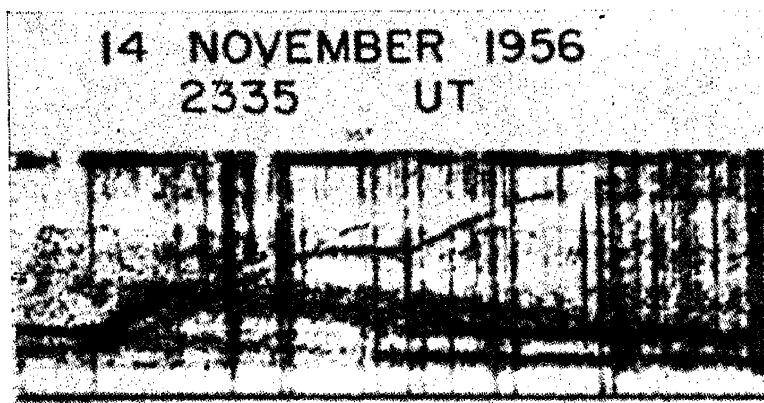
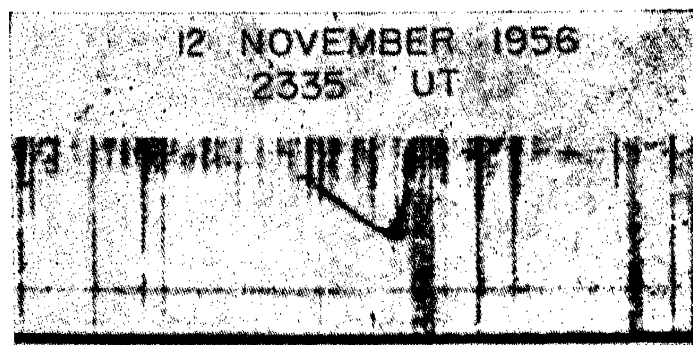
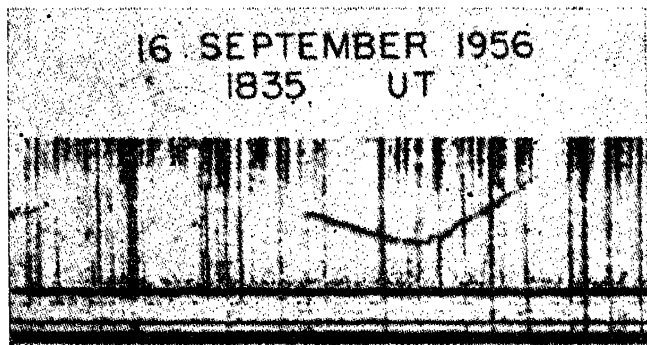
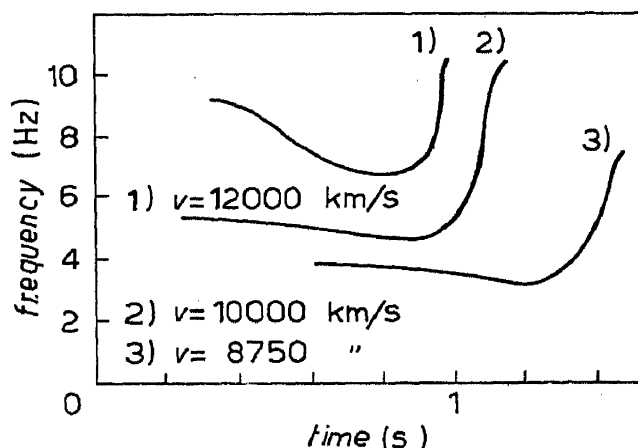


Fig. 3. — Samples of discrete emissions. After GALLET [7].

sions.) They have received suggestive names such as rising, falling tones, pseudo-nose whistlers, hooks and so on [7, 8] (Fig. 3 and 4).

The two kinds are mainly, but not exclusively, observed in high-latitude regions (50° to 65°), and are strongly correlated with magnetic disturbances [9-11], aurorae and even oxygen red lines air-glow [10, 12, 25].

Fig. 4. — Theoretical curves computed for the travelling wave mechanism. As the particle enters the ionosphere, f_0^2/f_H is no more constant and the emitted frequency increases very rapidly. After GALLET [7].



2.2. *Theoretical explanations.* — These emissions are undoubtedly produced by some interaction between particles of low energy ($V \sim (3\,000 \text{ to } 1\,000) \text{ km/s}$) with the exospheric medium. Once this emission has taken place, the electromagnetic waves are propagated with the whistler mode. It is known that this mode corresponds to large time-delay and great dispersion, but also to very well-defined ray-paths along the magnetic lines of force. Then from the knowledge of the whistler propagation, and from experimental frequency-time curves, one may expect to obtain results about energy and flux of the incoming particles, provided the actual mechanism of interaction is known.

2.3. *Processes related with the particle acceleration.* — Energy is radiated by a charged particle when it is accelerated. Apart from the «bremsstrahlung» which is negligible in this range of energy and frequency, we may expect a cyclotron or a synchrotron radiation.

Cyclotron radiation of protons has been proposed to explain the noise in the 750 Hz band, and even some discrete emissions [13-15]. One assumes that protons radiate at their gyrofrequency f_i , but that, due to the Doppler shift, the frequency which is in fact emitted is given by

$$f = f_i / (1 - \beta n),$$

where $\beta = V/c$ and n is the phase refraction index for the emitted frequency f . Due to the great value of n , one obtains, even for low-velocity particles, a frequency f which is so large that—at least in numerical application, and except in the lower ionosphere—it has no more relation with the gyrofrequency f_i , which could have been replaced by everything.

Cyclotron radiation of electrons would be more efficient, the radiated power

being inversely proportional to the mass. But the frequency produced would be the electronic gyrofrequency f_H and this frequency cannot be propagated in the whistler mode. But relativistic electrons can radiate at a much lower frequency

$$(1) \quad f = f_H(1 - \beta^2)^{\frac{1}{2}}$$

and at some harmonics which can be propagated. We have a synchrotron radiation which may be involved only in rare events (the minimum energy required is of the order of 0.5 to 1 MeV), but which can give a detectable power [16].

We will not discuss the detail of these mechanisms but we will recall that due to the large refraction index, these kinds of emissions are greatly affected by Doppler shift, as it has been already shown. Moreover, we need a theory of emission by accelerated particles in a highly anisotropic medium [17, 18]; we do not have such a theory in the whistler case.

2.4. *Processes related with the particle velocity.* — There is another kind of emission, which is not related with some acceleration process, but which occurs when the velocity of the particle is greater than, or equal to the phase velocity of the electromagnetic wave.

The later case is similar to the travelling wave amplification mechanism, and was first suggested by GALLET [7, 19]. Writing that the particle velocity V (which is assumed to be parallel to the magnetic field) is equal to the phase velocity of the wave v_ϕ (which is a function of the frequency (*)), GALLET obtains an equation defining the emitted frequency. With some realistic assumptions, this equation possesses two roots. One of them is nearly equal to the gyrofrequency and cannot be propagated. The second one is given by

$$(2) \quad f = \beta^2 f_0^2 / f_H,$$

where f_0 is the plasma frequency.

Assuming that V and f_0^2/f_H are constant (the later assumption is nearly verified between 1500 km and 15 000 km) the emitted frequency is constant, and this mechanism may account for some discrete emissions (see Fig. 4). If a plausible value for the ratio f_0^2/f_H (~ 4 MHz) is introduced, one can deduce from the observed frequency (~ 5 kHz) a value of β and thus of the velocity V ($\sim 10\,000$ km/s).

(*) He considers only the strictly longitudinal case.

A more complete theory of the travelling-wave amplification in the exosphere has been developed by BRICE [20].

3. - Čerenkov effect in the whistler mode.

The more general case $V > v_\phi$ is that of the Čerenkov effect. That this was a possible mechanism of V.L.F. emission was first suggested by ELLIS [21]. He computed the power which was to be expected from this process and found that, with reasonable assumptions, it could be detected. It is of interest to give some details on this effect, because it seems able to explain a great part of the broad-band noise. It is possible to give a simple ray theory when the velocity of the particle is parallel to the magnetic field.

3.1. *Equations.* - Čerenkov effect generates a frequency f in the direction θ when

$$v_\phi(f, \theta) = V \cos \theta$$

or

$$(3) \quad \beta n(f, \theta) \cos \theta = 1.$$

The refraction index n will be given for convenience in dimensionless units (*)

$$(4) \quad n^2 = 1 + \frac{\alpha^2}{x(\cos \theta - x)} \approx \frac{\alpha^2}{x(\cos \theta - x)},$$

where

$$\alpha = f_0/f_H \quad \text{is characteristic of the medium,}$$

$$x = f/f_H \quad \text{is characteristic of the emitted frequency.}$$

Substitution from (4) into (3) gives:

$$(5) \quad \beta^2 \alpha^2 \cos^2 \theta - x \cos \theta + x^2 = 0$$

which shows that, for a given angle, we obtain two emitted frequencies, and that a given frequency is emitted in two different directions.

One can discuss the existence of the roots of such an equation but it is better to draw a geometrical picture of the phenomena.

(*) We assume that $-\pi/2 < \theta < +\pi/2$. Otherwise, another formula would hold, but the results would be the same.

3'2. *Geometrical considerations.* — We define the velocity curve φ_x as the locus of the points of polar co-ordinates v_φ and θ ; and the phase index curve Γ_x as the locus of the points of polar co-ordinates n , θ . (The curves being symmetrical with respect to the magnetic field, only the case $\theta > 0$ will be considered.) This is drawn on Fig. 5. The angle θ_0 is defined by

$$\cos \theta_0 = x.$$

For greater values of θ , there is no propagation (see eq. (4)).

Assuming that we have a particle of velocity V directed along the magnetic field, emission angles θ , for the frequency $f = x f_H$, are given by the intersection of the circle of diameter OV and of the curve φ_x .

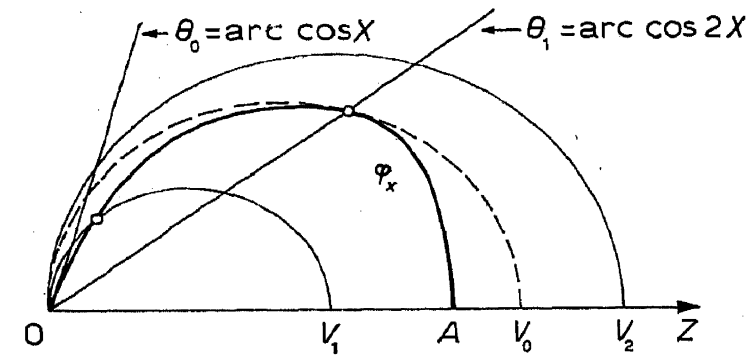


Fig. 5. — Čerenkov effect in an anisotropic medium. Z axis is the direction of the magnetic field and of the particle's velocity. Four different cases are possible; see text.

When V is too small ($V = V_1$), we obtain but one angle, which is not very different from θ_0 . This angle does not provide a physical solution because the effect of the collisions, which has been neglected, modifies the curve φ_x in the region involved. The guiding is poor and the absorption is strong.

When $|OV| = |OA|$, one angle is zero and we obtain the mechanism proposed by GALLET.

If V is too great ($V = V_2$ for example), there is no emission at all. This is different from what happens in the usual Čerenkov effect. *For too great values of the particle velocity, there is no emission.*

If $V = V_0$, the two angles join together and the circle is tangential to the φ_x curve. But the striking feature is that *the value V_0 does not depend on the frequency.* In other words, all φ_x curves are tangential to the same circle of diameter $|OV_0|$. The value of V_0 is

$$(6) \quad V_0 = c/2a = cf_H/2f_0.$$

This is due to the particular form of the eq. (4) [22].

Now we must recall some properties of the ray geometry in an anisotropic medium. If OMP is the wave normal direction (Fig. 6), it is known that the ray is propagated along the direction OT (perpendicular to the tangent PT), and that the ray velocity W is such that its projection on the wave normal is equal to the group velocity v_g .

Now, if all the φ_x curves are tangential to a circle, all the Γ_x curves are

tangential to a straight line perpendicular to the Z axis. This is related to the fact that the product $|OM| \cdot |OP|$ is a constant equal to c . Thus, all the frequencies emitted by Čerenkov effect when $V = V_0$ are propagated *exactly along the magnetic field*. Moreover, at the tangential point of the Γ_x curves with their envelope

$OP = n$

$OM = v_{\varphi}$

$OP = c$

$$\partial n / \partial x = 0$$

so that

$$(7) \quad v_g = v_\omega$$

and the velocity W of the ray is equal to V_0 , that is to say, to the velocity of the particle. This is true for all frequencies, and thus we have the following property:

All frequencies emitted from a particle of velocity V_0 by Čerenkov effect, are propagated in the same direction and with the same velocity [23, 24].

3.3. *Application.* — It is obvious that this must give a strong signal, because one receives, at the same time, all the power emitted in all frequencies. When there is a continuous flux of particles, one receives a continuous broadband noise.

There is a high-frequency cut-off, due to the fact that the angle θ_1 (see Fig. 5) is given by

$$(8) \quad \cos \theta_1 = 2x$$

and thus the allowed frequency range of emission when $V = V_0$ is reduced to

$$(9) \quad 0 < f < f_H/2.$$

Such broad-band noise with sharp high-frequency cut-off near 5 kHz are often recorded (see Fig. 1). If the mechanism we described is the correct one, the gyrofrequency at the point of emission ought to be 10 kHz. Taking the same value for f_0^2/f_{π} as before (4 MHz), the plasma frequency in the interaction region ought to be 200 kHz, and the velocity V , as given by (6), ought to be of the order of 8 000 km/s.

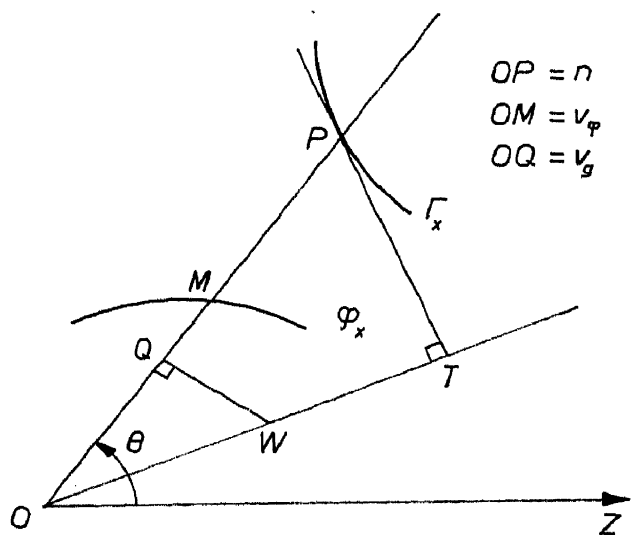


Fig. 6. - Direction and velocity W of the ray in an anisotropic medium. Γ_x and φ_x are the curves of index and of phase velocity at a given frequency $f = xf_H$ and at different angle θ .

3'4. *Emitted power.* – If the usual formula is valid in this case, the energy d^2E lost by a particle during a path of length dl in a bandwidth df is given by

$$(10) \quad \frac{d^2E}{dl df} = \frac{\pi e^2}{\epsilon_0 c^2} \cdot \left(1 - \frac{1}{\beta^2 n^2}\right) \cdot f,$$

where e is the charge of the particle, ϵ_0 the electric permittivity of vacuum. If we consider a cloud of particles of length l , of cross-section S , and of density N , the flux of power by unit bandwidth will be

$$(11) \quad \frac{d\Phi}{df} = \frac{\pi e^2}{\epsilon_0 c} \cdot \left(1 - \frac{1}{\beta^2 n^2}\right) V N \frac{l}{\lambda},$$

λ being the wavelength in vacuo [21]. In MKSA we have

$$d\Phi/df = 3.03 \cdot 10^{-35} (1 - 1/\beta^2 n^2) V N l / \lambda \quad \text{W} \cdot \text{m}^{-2} \cdot (\text{Hz})^{-1}.$$

Because of the perfect guiding when $V = V_0$, this flux is (in an ideal case) the flux which may be received at the earth. Moreover $\beta = 1/2a$, and the eqs. (4) and (8) give $n = a/x$ so that

$$1 - 1/\beta^2 n^2 = 1 - 4x^2,$$

which is independent of the plasma frequency. The value of x being between zero and one half, we do not encounter the difficulties, appearing in other media, the coefficient $(1 - 1/\beta^2 n^2)$ being always positive in our case.

If we assume that the length of interaction is 2 000 km, and that the density is about 10 e/cm³, we obtain for a frequency $f = 5$ kHz, and a velocity $V \sim 10$ 000 km/s, the following flux

$$10^{-19} \text{ W} \cdot \text{m}^{-2} \cdot (\text{Hz})^{-1}.$$

which is detectable, but somewhat lower than often observed (see concluding remarks).

3'5. *Limits of the theory.* – This theory of Čerenkov mechanism is of course, very rough. For example, we have not found any explanation of the fact that the high frequency cut-off seems to be greater at high latitude [25, 26].

We have assumed an homogeneous medium, and the question arises of what would become the theory in an actual exosphere. But it is valid for laboratory plasmas and may have some application in this field.

But, over all, a wave theory must be made, with the complete set of Max-

well's equations, as it has been already done for the ordinary waves generated by Čerenkov effect in a plasma [27].

4. - Conclusion:

As a conclusion, we can say that, if, theoretically, one can deduce from the noise recorded in low-frequency range, some data about the particles producing it, the method is, in some manner, very difficult. Experimental difficulties prevent us from having the actual line of force followed by the particle and the wave, the polarization and a good resolution time.

There are still some theoretical problems such as the general laws of emission of a charged particle in a highly anisotropic and dispersive medium. It seems probable that such a theory would mix together acceleration and velocity processes, and give a coherent picture of a great part of V.L.F. emission, when considering the actual velocity of particles, which is not strictly parallel or strictly perpendicular to the field. Also, important is the fact that when these emissions are produced by particles bunched closely (on distance much smaller than the wave length in the medium) there must be some coherence and the power emitted must be enhanced [28, 29].

It is also remarkable that all the suggested mechanisms—except the synchrotron radiation—involve particles whose velocities are always near to the value 10 000 km/s and thus we must have an explanation of the origin of such particles.

REFERENCES

- [1] G. R. A. ELLIS and D. G. CARTWRIGHT: *Nature*, **184**, 1307 (1959).
- [2] G. R. A. ELLIS, D. G. CARTWRIGHT and J. R. V. GROVES: *Nature*, **184**, 1391 (1959).
- [3] G. R. A. ELLIS: *Journ. Geoph. Res.*, **65**, 839 (1960).
- [4] J. Y. CROUCHLEY and N. M. BRICE: *Planet. Space Sci.*, **2**, 238 (1960).
- [5] J. AARONS, G. GUSTAFSSON and A. EGELAND: *Nature*, **185**, 148 (1960).
- [6] G. GUSTAFSSON and A. EGELAND: *Journ. Geoph. Res.*, **65**, 2749 (1960).
- [7] R. M. GAILLET: *Proc. I.R.E.*, **47**, 211 (1959).
- [8] J. H. POPE and W. H. CAMPBELL: *Journ. Geoph. Res.*, **65**, 2543 (1960).
- [9] G. R. A. ELLIS: *Low frequency electromagnetic radiation associated with magnetic disturbances*, in *AGARD Symposium on the Upper Atmosphere* (Paris, Mai 1959).
- [10] G. R. A. ELLIS: *Nature*, **186**, 229 (1960).
- [11] G. R. A. ELLIS: *Journ. Geoph. Res.*, **61**, 1705 (1960).
- [12] G. R. A. ELLIS: *Planet. Space Sci.*, **1**, 253 (1959).
- [13] J. W. MCARTHUR: *Phys. Rev. Lett.*, **2**, 491 (1959).
- [14] W. B. MURCRAY and J. H. POPE: *Phys. Rev. Lett.*, **4**, 5 (1960).
- [15] W. B. MURCRAY and J. H. POPE: *Journ. Geoph. Res.*, **65**, 3569 (1960).

- [16] R. GENDRIN: Note Technique du Centre National d'Etudes des Télécommunications PS 4102 (1961), p. 120.
- [17] R. Q. TWISS and J. A. ROBERTS: *Austral. Journ. Phys.*, **11**, 424 (1958).
- [18] V. IA. EIDMAN: *Sov. Phys. JETP*, **7** (34), 91 (1958).
- [19] R. M. GALLET and R. A. HELLIWELL: *Journ. of Res. of the N.B.S.*, **63D**, 21 (1959).
- [20] N. M. BRICE: *Journ. Geoph. Res.*, **65**, 3840 (1960).
- [21] G. R. A. ELLIS: *Journ. Atm. Terr. Phys.*, **10**, 302 (1957).
- [22] R. GENDRIN: *Compt. Rend. Acad. Sci.*, **251**, 1085 (1960).
- [23] R. GENDRIN: *Compt. Rend. Acad. Sci.*, **251**, 1122 (1960).
- [24] R. GENDRIN: *Planet. Space Sci.* **5**, 274 (1961).
- [25] L. H. MARTIN, R. A. HELLIWELL and K. R. MARKS: *Nature*, **187**, 751 (1960).
- [26] L. H. MARTIN: *Nature*, **187**, 1018 (1960).
- [27] M. H. COHEN: *Phys. Rev.* **123** 711 (1961).
- [28] J. V. JELLEY: *Čerenkov Radiation and Its Application* (New York, 1958), p. 45.
- [29] W. B. MURCRAY and J. H. POPE: *Proc. I.R.E.*, **49**, 811 (1961).

The Primary Cosmic Radiation.

C. J. WADDINGTON (*)

H. H. Wills Physical Laboratory, University of Bristol - Bristol

1. — Introduction.

The *primary* cosmic radiation may be defined as being a corpuscular radiation of high-energy which impinges on the earth's atmosphere, provided the term high energy is assumed to imply energies appreciably above conceivable thermal energies. This definition is convenient experimentally, but could be seriously misleading unless the importance of perturbations introduced by solar and terrestrial magnetic fields are taken into account. It is readily demonstrated that cosmic ray activity observed on or near the earth is closely related to solar phenomena such as flares or the sunspot cycle, but it is generally believed that, except during certain special and rather well defined periods, the majority of the primary particles do not originate in the sun (see Section 8). Consequently any discussion of the origin or propagation of these particles presupposes a knowledge of the properties of the radiation before it experiences any solar or terrestrial influences. The properties of the « *galactic* » (**) radiation must therefore be inferred from those of the primary radiation, and in some cases this can lead to serious uncertainties, particularly since much of the most experimentally-reliable data presently available were obtained during the recent period of high solar activity.

This definition of the primary radiation as being corpuscular in nature excludes any possible high-energy γ -rays, even though such quanta could well be produced by the cosmic radiation passing through interstellar matter. The study of these quanta forms the separate and exciting new discipline of γ -ray astronomy.

(*) Present address: School of Physics, University of Minnesota, Minneapolis, Minn.

(**) Here the term « galactic » is not intended to exclude any extra-galactic radiation from consideration.

Any theory or discussion of the origin of these energetic particles must be related to experimental data about the galactic radiation. The most important of these data are:

- i) The charge distribution.
- ii) The energy or rigidity distribution.
- iii) The spatial distribution.

These first two are, of course, interdependent, since strictly the charge distribution can only be specified for particles at or above a given energy or rigidity.

Current knowledge regarding these distributions will be summarized in this article, together with an introduction to the significance of these results to theories of the origin of the cosmic radiation. In addition there is some discussion of the future experiments which should be undertaken into specific points of uncertainty.

2. - The chemical distribution.

2.1. *Experimental techniques.* - The two principal particle detectors which have been used to study the detailed features of the primary radiation are:

- a) The combination Čerenkov-scintillator (Č-S) array developed by McDONALD [1].
- b) Nuclear photographic emulsions.

The main features of the Č-S array are shown in Fig. 1, while Fig. 2 shows the responses of the two counters plotted as a function of particle velocity expressed as a fraction of the velocity of light, β . In both cases the outputs vary as the square of the nuclear charge, Z . The Č-S array has proved of great value in studying the hydrogen and helium nuclei in the primary radiation, but due partly to saturation problems and partly to the small geometrical collection factor it has not yet been systematically applied to investigations of heavier nuclei.

Emulsions, on the other hand, have been used mainly to study primary nuclei heavier than hydrogen, although recently hydrogen nuclei have also been the subject of experiments in these detectors (WADDINGTON [3]). Since much of the uncertainty in our knowledge of the abundances of the heavy nuclei is due to the limitations of emulsions as a particle detector, it is worth considering these limitations in slightly more detail.

Photomicrographs of the tracks produced by fast ($\beta \sim 1$) nuclei of different Z in emulsions having differing levels of development are shown in Plate I. At $\beta \sim 1$ the ionisation produced, I , is proportional to Z^2 . However, as can be seen, the cores of the tracks produced in the emulsion rapidly saturate

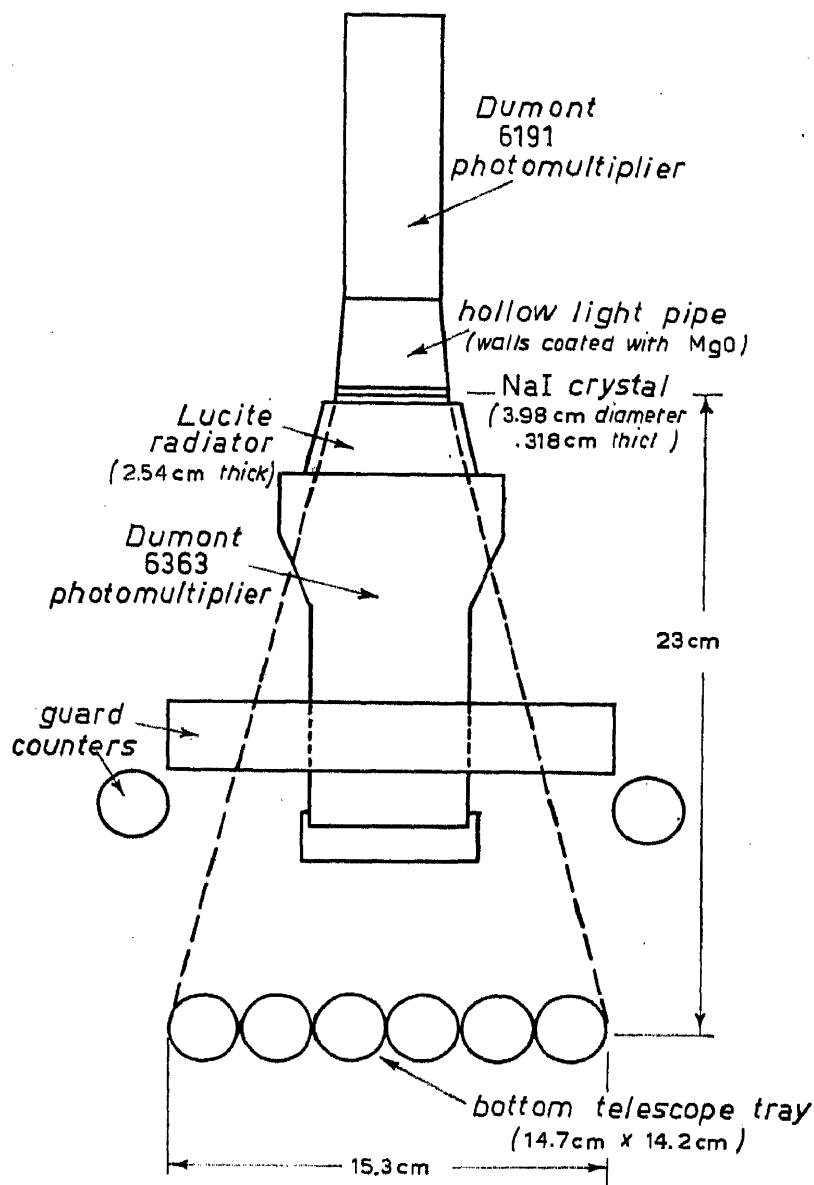


Fig. 1. - Outline drawing of McDonald's Č-S array (after McDONALD [1]).

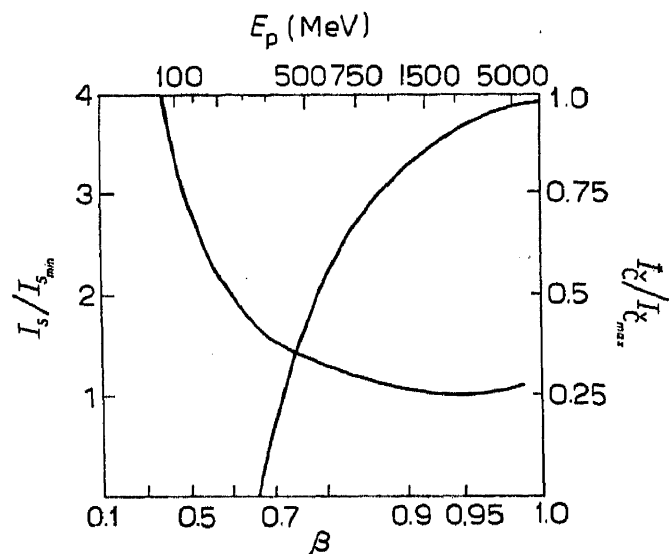


Fig. 2. - Č-S array output as a function of the velocity β and energy, E_p , of a proton. Čerenkov counter output I_c , scintillator counter output I_s . (After McDONALD and WEBBER [2]).

with increasing Z , and a determination of I generally involves a measurement of the δ -ray density. As is well known, the measurement of these densities is subjective if made by an observer, and sensitive to background effect if determined by densitometer techniques. Furthermore, for high Z values ($Z \geq 10$)

the percentage differences in I between individual charges, $\sim (200/Z)\%$, are small. Consequently charge discrimination, which demands measurements which are at least three times as precise as these percentage differences, *i.e.* $\pm(70/Z)\%$, has seldom, if ever, been obtained for $Z \geq 10$, although the results reported by KRISTIANSSON *et al.* [4] suggest that very painstaking densitometer techniques may be able to reach near to this limit.

2.2. Magnetic rigidity. - Because the motion of cosmic ray particles appears to be governed by magnetic fields it is found that although the energy or mo-

mentum of a particle is the quantity usually measured experimentally, it is often of more fundamental interest to consider its magnetic rigidity, R . This quantity R is defined in the relation for the radius of curvature, r , of a charged particle moving in a magnetic field, of:

$$(1) \quad r = \frac{1}{300} \cdot \frac{cp}{H_{\perp} Z} = \frac{1}{300} \cdot \left(\frac{Acp_N}{Z} \right) \frac{1}{H_{\perp}} = \frac{10^9}{300} \cdot \frac{R}{H_{\perp}} \text{ cm},$$

when R is expressed in GV; p , the momentum, or p_N , the momentum per nucleon, in eV/c; $H_{\perp}$, the magnetic field perpendicular to the line of flight of the particle, in gauss; and Z , the atomic number, in units of the electron charge. Fig. 3 shows R plotted against energy for electrons and protons, and against energy per nucleon for nuclei

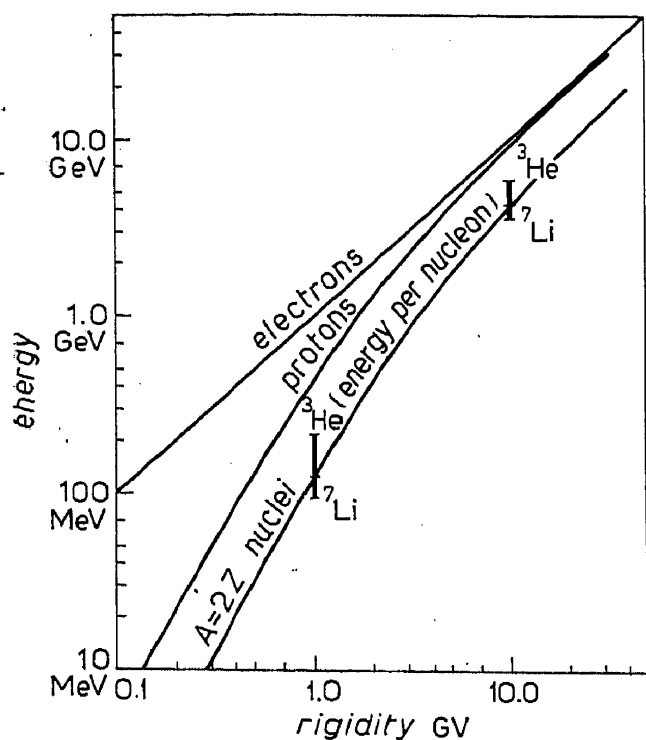


Fig. 3. — The variation of rigidity, in GV, with kinetic energy or energy per nucleon. Curves are shown for electrons, protons and nuclei having $A = 2Z$, while specific values are given for ${}^3\text{He}$ and ${}^7\text{Li}$ nuclei, as examples of nuclei with $A \neq 2Z$.

against energy per nucleon for nuclei

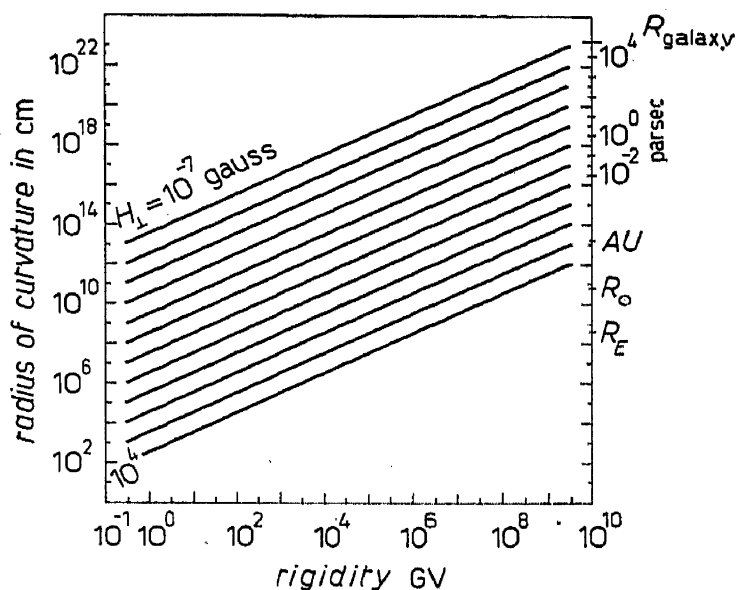


Fig. 4. — The variation of rigidity, in GV, with radius of curvature, in cm, for various values of the perpendicular magnetic field $H_{\perp}$, in gauss.

having $A = 2Z$, while Fig. 4 shows R plotted against r for various values of H .

The effect of the geomagnetic field at any one location and looking in any particular direction is to remove all those particles which have R less than some minimum value, the threshold rigidity. As a result measurements made at any point under the geomagnetic field can only result in the detection of galactic particles whose rigidity exceeds the threshold value appropriate to

that place. These threshold rigidities vary, for vertically incident particles, between about 16 GV near the equator, to zero at the magnetic poles.

2.3. *Solar modulation.* — Because it is often necessary to compare experimental results determined at different times and under different conditions of solar activity it is most important to have some way of estimating the magnitudes of solar-induced variations. Neutron monitors provide a continuous and sensitive record of the cosmic ray activity at sea level, and these data may often, though not always, be correlated with effects, such as intensity changes, observed above the atmosphere. In this paper the data from the Ottawa neutron monitor have been used as a measure of the cosmic ray activity, partly because this is a station at a relatively high geomagnetic latitude which has been operating since 1955, and partly simply because of the availability of the data (*).

The results are expressed in the natural logarithm form introduced by WADA [5] where

$$(2) \quad W_p = \ln(N_p/1000) - W_p^* ;$$

$W_p^* = 1500$ and N_p is the daily mean bi-hourly counting rate.

W_p is expressed in units of 0.1% . Fig. 5 shows the mean *monthly* values of W_p since 1955, from which it can be seen that as a first approximation the activity

can be regarded as two-valued. Before 1957, solar «minimum», W_p had an average value of ~ 330 , while from 1957 onwards, solar «maximum», $W_p \sim 200$. It would seem probable, from the characteristically rapid increase in solar activity, compared with the slow decrease which was observed during previous solar cycles, that it will not be possible to make such a clear distinction between solar maximum and the next period of solar minimum. However, it should be noticed that the correlation between the solar activity, as measured by the Zurich sunspot number, and the cosmic ray activity is not complete, and these modulation effects may show some kind of threshold.

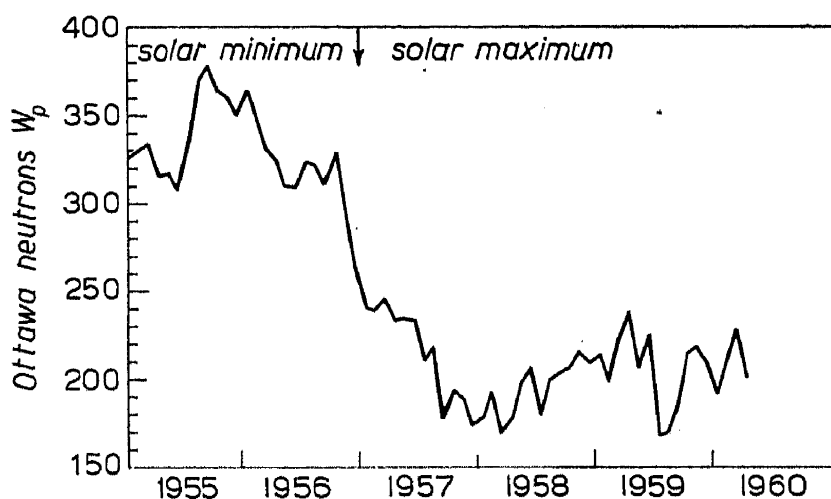


Fig. 5. — The mean monthly values of W_p from the Ottawa neutron monitor since the beginning of 1955.

(*) The assistance of Dr. D. C. ROSE in obtaining these data is acknowledged with gratitude.

3. - The chemical distribution at $R > 4.5$ GV.

At present the best data available on the chemical distribution are those obtained for particles having a rigidity greater than about 4.5 GV. Since only some 10% of these particles have $R > 22$ GV, these data essentially refer to the chemical distribution of particles having rigidities between these two values. At a rigidity of ~ 4.5 GV all the primary particles have energies so great that $\beta \sim 1$ and thus I is proportional to Z^2 , which produces a considerable simplification of the experimental problems of charge identification. At the same time this rigidity is not so high that it is unduly difficult to obtain results of reasonable statistical weight.

A geomagnetic vertical threshold rigidity of ~ 4.5 GV is observed along an «isocosm» (*) which passes through Texas and Northern Italy, and it is mainly at these places that the experiments considered in the following sections have been performed.

3.1. Hydrogen nuclei. - Both numerically and energetically these nuclei represent the most important single component of the primary radiation. That they are predominantly protons, rather than deuterons or tritons, is shown by the failure to observe any appreciable intensity of nonrelativistic singly charged particles at threshold rigidities where such multiply charged nuclei are observed. Nevertheless, for various experimental reasons the available information about these nuclei is somewhat less detailed and precise than that on the rarer heavier nuclei.

a) At this rigidity, these nuclei produce minimum ionization. Consequently (see Plate I) they are not readily detected in nuclear emulsions unless they produce some secondary effect, such as a nuclear interaction.

b) Unless the measurements are made on space probes far removed from the earth, which has not yet been done, these nuclei have to be distinguished from amongst a relatively copious background of albedo and other secondary particles which also produce minimum ionization. In general this imposes the necessity of at least determining the direction of flight of each particle.

c) There is also a re-entrant albedo of downward moving particles, presumably mainly electrons, with rigidities below the threshold value, which must either be discriminated against or corrected for.

(*) A line on the earth's surface along which the vertical threshold rigidity is constant.

For these reasons, until recently the only data on the intensities of these particles which were apparently reliable were those provided by the Č-S array, McDONALD [1], McDONALD and WEBBER [2]. Even these results were uncertain in so far as they depended on an extrapolation to the top of the atmosphere

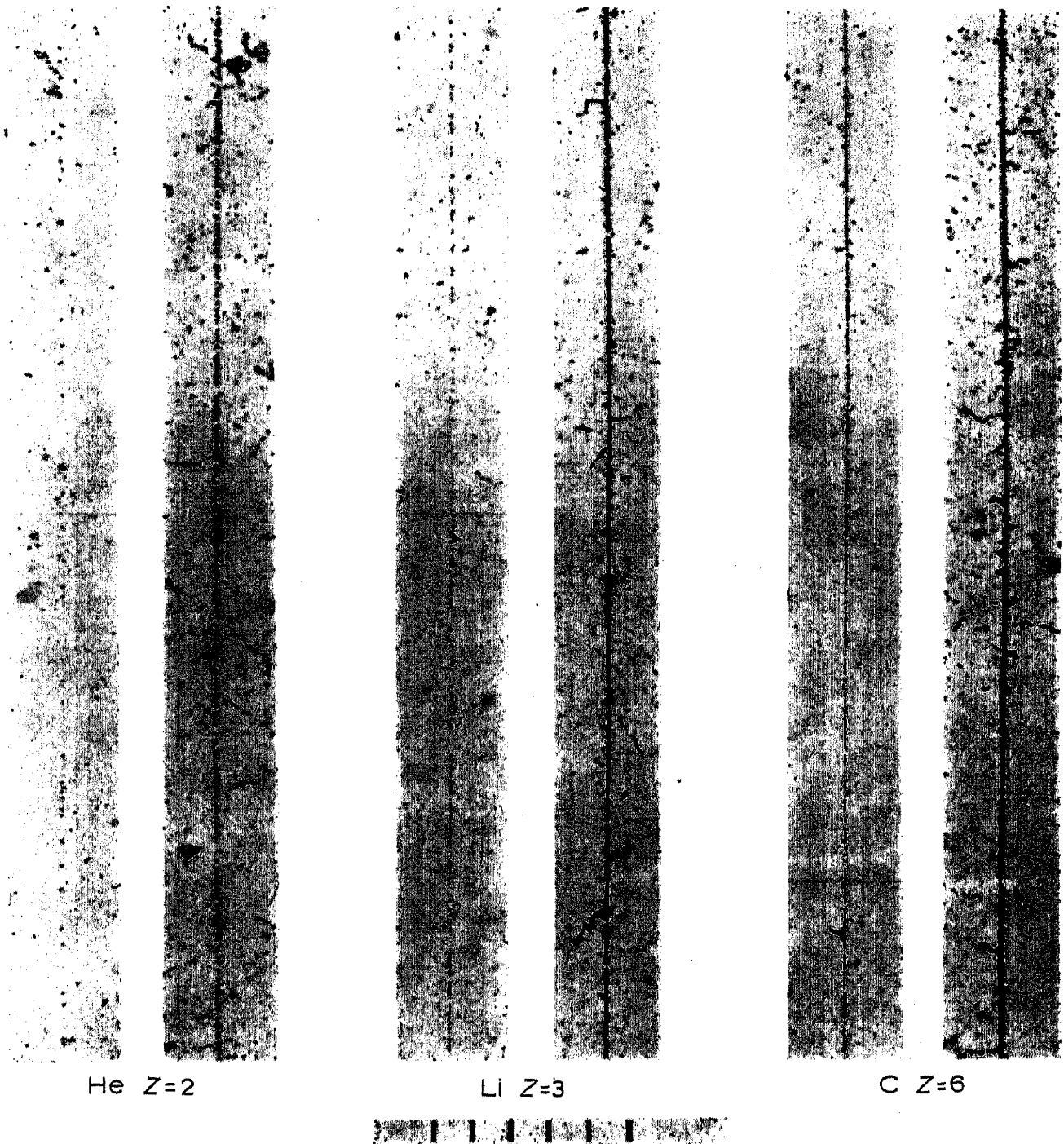


Plate I. — Photomicrographs of the tracks produced in Ilford G5 emulsion by fast, $\beta \approx 1$, nuclei. In each pair, the track on the left was detected in a severely underdeveloped emulsion, while that on the right was in a rather heavily developed emulsion. In spite of the greater background on the right-hand prints, they were actually flown for less than half the time of the others.

and an estimate of the—relatively small—correction for the re-entrant albedo.

Recent measurements in nuclear emulsion, WADDINGTON ([3] and unpublished), confirm the validity of the $\check{O}$ -S measurements at these rigidities and show that no *serious* errors can have been introduced by any of the assumptions or

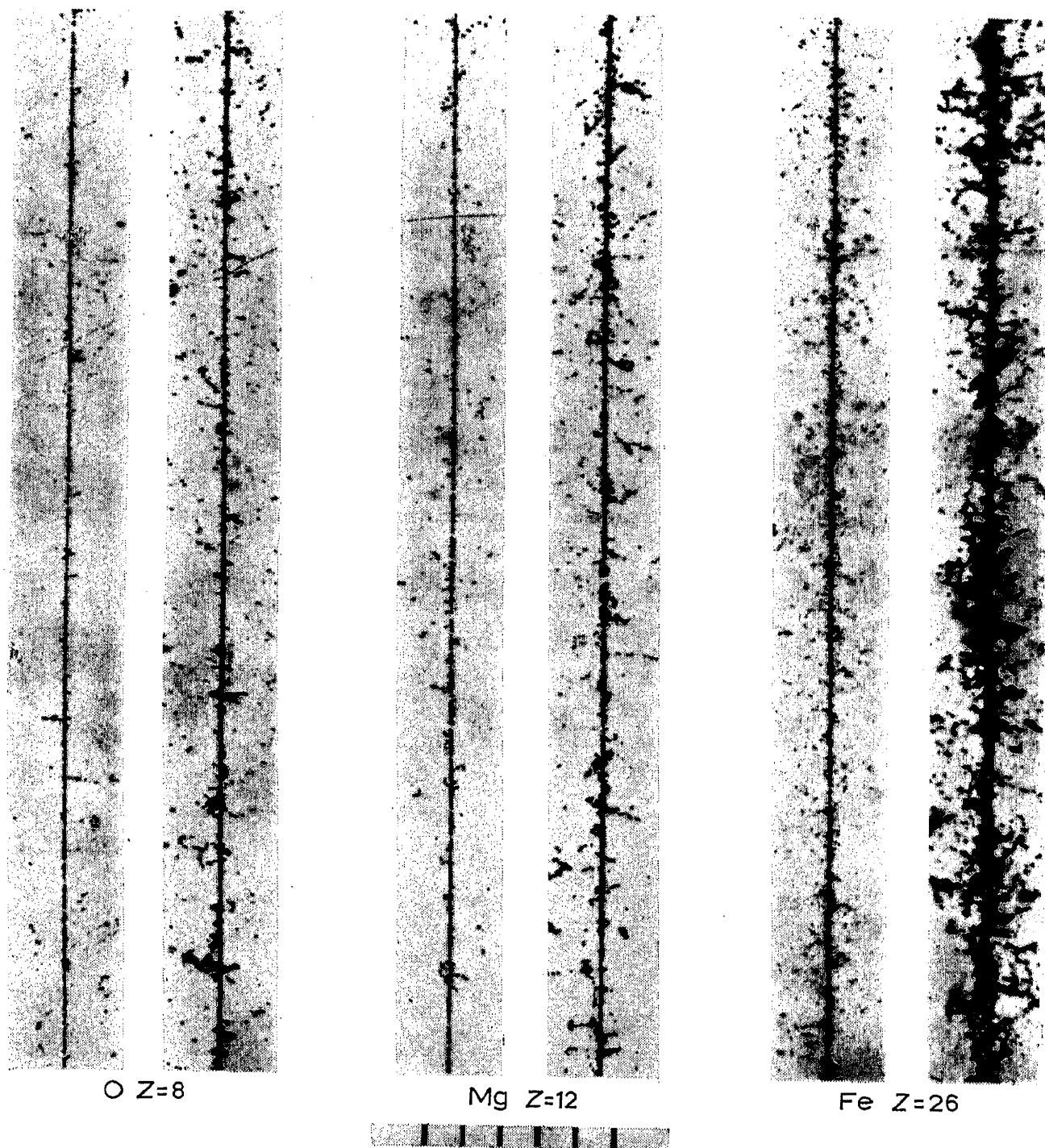


Plate I. - Photomicrographs of the tracks produced in Ilford G5 emulsion by fast, $\beta \approx 1$, nuclei. In each pair, the track on the left was detected in a severely underdeveloped emulsion, while that on the right was in a rather heavily developed emulsion. In spite of the greater background on the right-hand prints, they were actually flown for less than half the time of the others.

corrections made. No such check, however, exists for the Č-S measurements made at lower rigidities, nor were any measurements made prior to 1955. Fig. 6 shows the intensity values found by the Č-S array and emulsion at a threshold

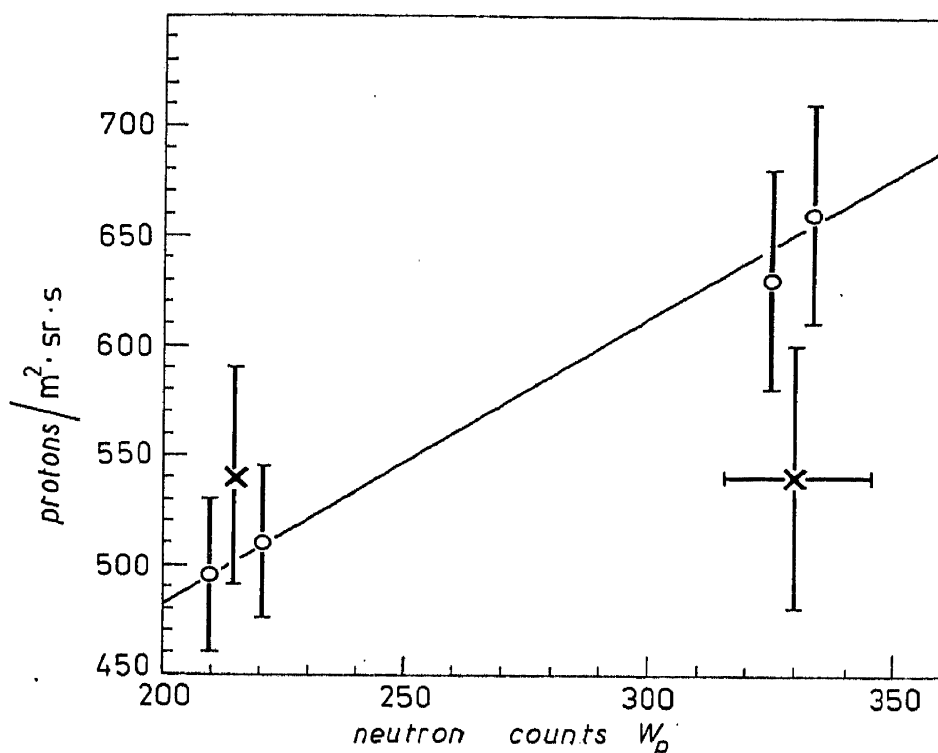


Fig. 6. — The proton intensities determined with the Č-S array, O, and with emulsions, X with $R > 4.5$ GV, as a function of W_p .

rigidity of 4.5 GV, plotted against the W_p values appropriate to the day of each experiment. The agreement between the Č-S and emulsion determinations is reasonable, particularly when the considerable extrapolations involved in the first emulsion determination, that made during solar minimum at a high W_p , is taken into account. From these results the best value for the intensity at solar minimum, $J_p^*(0)$, is (610 ± 30) protons/m²·sr·s while that at solar maximum, $J_p(0)$, is (515 ± 25) protons/m²·sr·s. This value of $J_p^*(0)$ may be compared with that of 616 ± 23 derived from semi-independent data in the next section.

The precise isotopic composition of the hydrogen nuclei is unknown, there being no experimental evidence regarding the magnitude of any possible deuteron component. DANIEL *et al.* [6] have suggested an experiment which should be capable of detecting a deuteron intensity $> 5\%$ of the hydrogen intensity. This limit may be compared with that calculated by HAYAKAWA *et al.* [7] of $\leq 0.3 \cdot X\%$, where X is the thickness of interstellar material traversed by the producing particles from the source to the earth. If the results described later which suggest that at least for some components X is ≥ 12 g/cm² are confirmed, then a finite deuteron intensity may be just measurable in this manner.

3.2. *Helium nuclei.* — The primary helium nuclei have been investigated by many different groups, using both Č-S arrays and emulsions as detectors. Since there is no albedo and these nuclei are easily detectable in emulsions, the intensity values are known with some precision. The mean intensity of α -particles having $R > 4.5$ GV during solar minimum has been calculated, (WADDINGTON [8]), from the results of a number of experiments as

$$J_{\alpha}^{*}(0) = (88 \pm 2) \text{ } \alpha\text{-particles/m}^2\cdot\text{sr}\cdot\text{s}.$$

These values were calculated from measurements made below the top of the atmosphere, but since the corrections due to absorption and production by heavier nuclei are relatively small and appear to be well known, any systematic errors should certainly be less than 5%.

WEBBER [9] has listed a considerable number of determinations of the ratio of the proton to α -particle intensities, $\Gamma_{p\alpha}(R)$, made with the Č-S array both at this rigidity and at others. Table I shows the mean values of $\Gamma_{p\alpha}(R)$ obtained from these data and the emulsion experiment at various rigidities. For $R < 4.5$ GV these results depend solely on the Č-S array, operated under conditions where the background conditions are most serious.

TABLE I. — *Values of $\Gamma_{p\alpha}(R)$ for various rigidities.*

Rigidity	> 1.0 GV	> 1.5 GV	> 2.5 GV	> 4.5 GV	> 17.0 GV
$\Gamma_{p\alpha}(R)$	7.1 ± 0.4	7.3 ± 0.3	6.2 ± 0.7	6.9 ± 0.5	7.4 ± 1.0

These values of $\Gamma_{p\alpha}(R)$ are consistent with a unique mean value of 7.0 ± 0.2 . If this mean is combined with the intensity value of $J_{\alpha}^{*}(0)$ given above, a value of $J_p^{*}(0) = (616 \pm 23)$ protons/m²·sr·s is obtained which is in good agreement with that given previously.

The ratio of the intensity of protons to that of α -particles above a given energy per nucleon, *i.e.* above a given velocity, $\Gamma_{p\alpha}(E)$, which is the quantity to be used in any comparison with the cosmic abundances (Section 3.4), can only be determined from these results if the energy spectra of one or both components are known. Unfortunately, a consequence of the constancy of $\Gamma_{p\alpha}(R)$ is that $\Gamma_{p\alpha}(E)$ varies except when $E \gg m_0c^2$, decreasing with decreasing E . In addition, it also depends on the isotopic composition of the two components, which are not known. The calculation of these ratios is discussed in Appendix I, from which it appears that when $E \gg m_0c^2$

$$\Gamma_{p\alpha}(E) = 19.8 \pm 1.2,$$

which is an *upper* limit to the possible value above any lower energy.

The ratio above a given total kinetic energy, $\Gamma_{\text{px}}(U)$, is derived from this value of $\Gamma_{\text{px}}(E)$, and, since it is only of interest when considering extensive air showers, when $E \gg m_0 c^2$, has an actual value of

$$\Gamma_{\text{px}}(U) = 2.48 \pm 0.15.$$

3.3. $Z \geq 3$ nuclei. — The relative abundances of these nuclei have been a source of controversy since their existence in the primary radiation was first demonstrated. Due originally to the subjective nature of emulsion techniques, this controversy has more recently centred on the magnitude of the corrections that should be made for the effects introduced by the overlying atmosphere. However, because of the considerable improvements in balloon flying techniques that there have been in recent years experiments have been made at progressively smaller depths in the atmosphere, thus reducing the uncertainties introduced by these corrections. Undoubtedly in the near future they will be still further reduced by experiments conducted in satellites. As a result it now appears that the relative abundances of these nuclei are known with a fair degree of certainty, although many of the finer details still need to be determined.

An analysis of the available experimental data is conveniently started by first considering the ratios between various charge groups, and then discussing the determination of the absolute intensities.

At present, results are available from three experiments which were performed within 5 g/cm² of the top of the atmosphere and which were intended specifically to investigate the abundance of the L -nuclei, $3 < Z < 5$. Of these, that due to FREIER *et al.* [10] is of relatively low statistical weight, while those of O'DELL, SHAPIRO and STILLER [11] and of VAN HEERDEN and JUDEK [12] both used emulsions exposed on the same balloon so that they were not entirely independent.

In what follows it is assumed that the relative abundances of these nuclei are independent of solar activity.

a) Relative abundances.

1) The ratio of L to S -nuclei, Γ_{LS} .

The values found for the ratios of L -nuclei to all heavier nuclei, S -nuclei, $\Gamma_{LS}(R, x)$ (*), are shown in Fig. 7 as a function of the depth of overlying atmos-

(*) Since $\Gamma_{LS}(R, x)$ is essentially equal to $\Gamma_{LS}(E, x)$, see eq. (14), though not to $\Gamma_{LS}(U, x)$, no distinction will be made between these two quantities.

phere, x , expressed in g/cm^2 . Also shown are a number of values obtained in earlier experiments which had reasonable statistical weight and $x \leq 12 \text{ g/cm}^2$. These results may be corrected to the top of the atmosphere either i) by making a calculation for the absorption and production of nuclei in the over-

lying atmosphere; the diffusion correction; or ii) by using the experimental values of $\Gamma_{LS}(x)$ to construct a growth curve.

The first of these methods depends on a knowledge of the various appropriate parameters describing the absorption and production of nuclei in the atmosphere. In general these have not been determined directly in air, but have had to be inferred from measurements made in emulsions by using assumptions of questionable validity. The second method makes the dubious assumption that the results used are a homogeneous sample unaffected by systematic errors. By excluding those earlier measurements made with $x > 12 \text{ g/cm}^2$, together with that of APPA

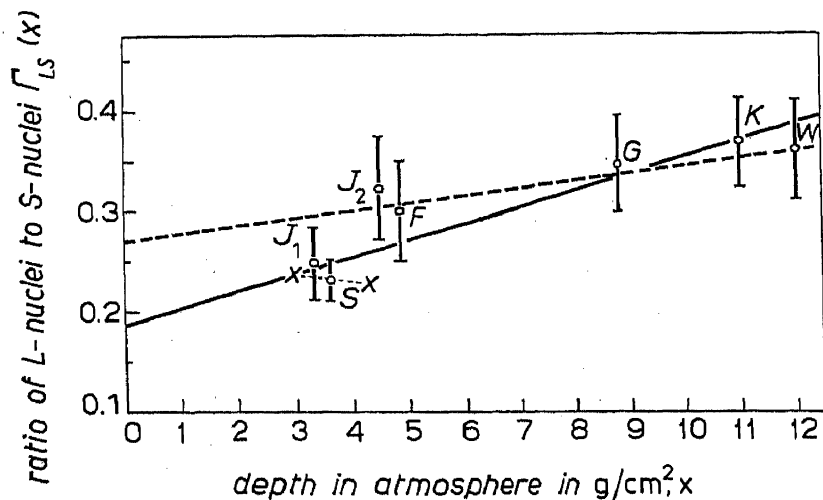


Fig. 7. - Ratio of L -nuclei to S -nuclei, $\Gamma_{LS}(x)$, plotted as a function of x , the depth in the atmosphere, in g/cm^2 . Points are: W , WADDINGTON [13]; K , KOSHIBA *et al.* [14]; G , GARELLI *et al.* [15]; F , FREIER *et al.* [10]; J , VAN HEERDEN and JUDEK [12], J_1 , $\vartheta \leq 30^\circ$, J_2 , $30^\circ < \vartheta \leq 60^\circ$; and S , O'DELL *et al.* [11]; crosses show $\vartheta \leq 30^\circ$, $30^\circ < \vartheta \leq 60^\circ$. Dashed straight line shows diffusion extrapolation to top of atmosphere normalized to point W , while solid straight line is the least squares fit.

RAO *et al.* [16], many of the more grossly incompatible results have been eliminated, but the difficulty of estimating the reliability of the data used is still present.

The use of the diffusion correction is shown in Fig. 7, where the calculated growth curve, assumed for simplicity to be linear, is normalized to the point at $x = 12 \text{ g/cm}^2$, giving $\Gamma_{LS}(0) = 0.27$. This curve is calculated from the values of the diffusion parameters published previously (WADDINGTON [8]), which were derived from emulsion data. Experiments of low statistical weight using emulsions interleaved with materials containing light elements suggest values for these parameters which would lead to a somewhat lower value of $\Gamma_{LS}(0)$, so that it is probably reasonable to take 0.27 as an upper limit. Notice also that the curve shown is arbitrarily normalized, and can be displaced along the $\Gamma_{LS}(x)$ axis.

Using the method of least squares, and weighting each point in proportion to its statistical weight, the best linear fit to the data is

given by

$$(3) \quad \Gamma_{LS}(x) = (0.0166 \pm 0.0034)x + (0.186 \pm 0.021) \cdot (*) .$$

It may be noticed from Fig. 7 that the difference between this curve and that obtained from the diffusion correction is almost entirely due to the results of O'DELL *et al.* [11]. It is thus certainly not unreasonable to take $\Gamma_{LS}(0) = 0.16$ as a lower limit.

Therefore a best value of $\Gamma_{LS}(0)$ may be taken as

$$\Gamma_{LS}(0) = 0.21 \pm 0.05 .$$

A more precise value will only come from measurements made at x less than 1.0 g/cm^2 , presumably from space vehicles.

2) The ratio of H to M -nuclei, Γ_{HM} .

The long standing controversy about the abundance of L -nuclei has tended to divert attention from the appreciable uncertainties that still exist in our knowledge of the abundances of the heavier nuclei. These are due to:

- i) The experimental difficulty of obtaining charge resolution between nuclei with $Z > 8$.
- ii) The increased importance of the corrections introduced by the overlying atmosphere, due to the larger interaction cross-sections of heavier nuclei.
- iii) The low abundances of some of the heavier nuclei which makes it difficult to obtain results of reasonable statistical weight.

It is usual to separate the S -nuclei into M -nuclei, $6 < Z < 9$, and H -nuclei, $Z \geq 10$; although because of the wide range of charges in this latter group, from 10 to at least 26, it, in turn, is frequently subdivided into two or more groups. As a result of the apparent low abundance of nuclei with $15 < Z < 19$ this separation is usually based on a division in this range of charges. Here, nuclei with $10 < Z < 19$ will be denoted as $\mathcal{H}$ -nuclei, those with $20 < Z < 28$ as $\mathcal{VH}$ -nuclei.

The ratio of H to M -nuclei, $\Gamma_{HM}(0)$, may be determined by similar procedures to those used to determine $\Gamma_{LS}(0)$. Fig. 8 shows the experimental values of $\Gamma_{HM}(x)$, obtained in the same experiments used to determine $\Gamma_{LS}(0)$, plotted as a function of x . As in Fig. 7 the figure shows the diffusion curve normalized

(*) These errors are calculated from residuals, and are thus, due to the small number of points, subject to considerable uncertainties, nor, expressed in this form, are the coefficients statistically independent.

to the point at $x=12$ g/cm² (*). This approach leads to $\Gamma_{HM}(0) = 0.38$. Also shown is the best least squares fit, given by

$$(4) \quad \Gamma_{HM}(x) = (0.00143 \pm 0.00035)x + (0.318 \pm 0.003).$$

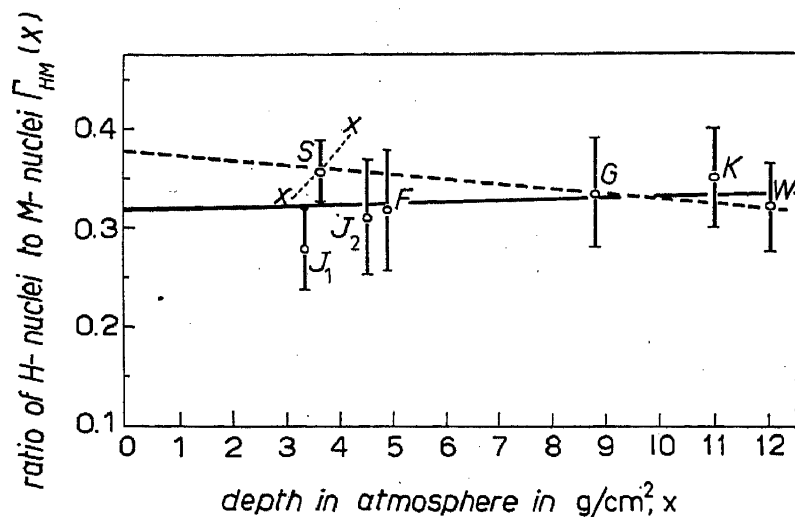


Fig. 8. — Ratio of H -nuclei to M -nuclei, $\Gamma_{HM}(x)$, plotted as a function of x , the depth in the atmosphere. Experimental points and lines have similar significance as in Fig. 7.

growth curve which predicts $\Gamma_{HM}(0) = 0.31 \pm 0.02$.

It appears reasonable therefore, to assume that $\Gamma_{HM}(0)$ lies between 0.38 and 0.30, leading to a best value of

$$\Gamma_{HM}(0) = 0.34 \pm 0.04.$$

This value also, will only be improved on by measurements made at very small values of x .

3) The ratio of $\mathcal{VH}$ to $\mathcal{H}$ -nuclei, $\Gamma_{\mathcal{VH.H}}$.

Unfortunately none of those experiments done at x less than 5 g/cm² have attempted to resolve the H -nuclei, and so the best value for $\Gamma_{\mathcal{VH.H}}(0)$ is still that quoted previously (WADDINGTON [8]), from an analysis of data at x greater than 5 g/cm², which was extrapolated to the top of the atmosphere by using the diffusion corrections.

(*) This value is in fact a little lower than that quoted previously since the earlier value appears to have been affected by an unfortunate arithmetical error. The value of 0.320 ± 0.045 quoted here is obtained from the values published by WADDINGTON [13] of $J_M = 4.29 \pm 0.33$ nuclei/m² sr s and $J_H = 1.35 \pm 0.16$ nuclei/m² sr s after being corrected through 4.8 g/cm² of emulsion; for which the corrections are well known.

Therefore

$$\Gamma_{\nu\mathcal{H}\mathcal{H}}(0) = 0.37 \pm 0.08.$$

This value is particularly sensitive to the assumed diffusion parameters and urgently needs to be verified by experiments at small depths in the atmosphere.

4) The ratios of individual nuclei to all nuclei, Γ_{iX} .

It does appear that it is just reasonable to attempt to determine the abundance of individual L and M -nuclei. If $Z \geq 3$ nuclei are denoted as X -nuclei, then $\Gamma_{Xi}(0)$ may be determined by procedures similar to those used to determine $\Gamma_{Ls}(0)$ and $\Gamma_{HM}(0)$. For these nuclei, however, the interaction parameters are known with even less precision than they were for the charge groups and

TABLE II. - Values of $\Gamma_{iX}(x)$ expressed as percentages.

Ref.	x g/cm ²	Li	Be	B	C	N	O	F	$Z \geq 10$
WADDINGTON [13]	13	7.8 ± 0.8	6.7 ± 0.7	13.1 ± 1.3	23.8 ± 2.4	14.1 ± 1.4	13.1 ± 1.3	4.1 ± 0.4	17.3 ± 1
GARELLI <i>et al.</i> [15]	8.8	5.7 ± 1.7	7.5 ± 2.1	12.5 ± 2.7	23.4 ± 2.4 (*)	15.0 ± 2.2 (*)	15.4 ± 2.2 (*)	1.9 ± 1.0	18.6 ± 4
FREIER <i>et al.</i> [10]	5	5.2 ± 1.6	4.7 ± 1.5	14.4 ± 2.5	—	59.6 ± 7.0	—	—	18.1 ± 3
VAN HEERDEN and JUDEK [12]	4	4.1 ± 0.8	3.4 ± 0.7	15.0 ± 1.6	19.1 ± 1.8	19.9 ± 1.8	12.8 ± 1.4	5.4 ± 0.9	20.3 ± 1
O'DELL <i>et al.</i> [11]	4	5.9 ± 0.9	2.9 ± 0.6	9.6 ± 1.1	29.4 ± 1.9	9.5 ± 1.1	19.0 ± 1.6	2.3 ± 0.5	21.4 ± 1

(*) Assumes same percentage total of CNO nuclei as GARELLI *et al.*, but uses ORSINI's [18] relative abundances.

attempts to fit least square relations are affected by greater statistical errors. Table II shows the values of Γ_{iX} , expressed as percentages, obtained in several experiments, as well as the mean values obtained previously (WADDINGTON [8]), from earlier experiments. Table III shows the values of the constants a and b calculated by the method of least squares in the relations

$$\Gamma_{iX} = ax + b$$

together with the resulting values of $\Gamma_{iX}(0)$, expressed as percentages. Also shown for comparison are the values suggested by WADDINGTON [8] from a diffusion correction to the mean values quoted in Table II. The agreement

is reasonable and it may be noticed that the values of $\Gamma_{LS}(0)$ and $\Gamma_{HM}(0)$ which can be derived from this table are not greatly different from those values assumed above to be the best available.

TABLE III. - Values of a and b and a comparison of $\Gamma_{iX}(0)$ values.

	a	b	b as %	$\Gamma_{iX}(0)$ as % from WADDINGTON
Li	0.32 ± 0.10	3.57 ± 0.80	3.9	5.2
Be	0.403 ± 0.076	1.63 ± 0.59	1.7	4.3
B	0.19 ± 0.28	10.7 ± 2.2	11.6	11.9
C	$-(0.036 \pm 0.80)$	24.0 ± 6.0	26.0	25.1
N	0.22 ± 0.66	11.5 ± 5.3	12.4	14.9
O	$-(0.259 \pm 0.385)$	16.6 ± 3.4	17.9	14.5
F	0.12 ± 0.16	2.4 ± 1.6	2.6	4.0
$Z \geq 10$	$-(0.387 \pm 0.113)$	22.1 ± 0.9	23.9	20.1

This analysis assumes that there are no systematic errors in the data used. However, an examination of the individual Γ_{iX} values shows evidence for serious discrepancies, *e.g.* $\Gamma_{\alpha X}$ and Γ_{NX} as observed by O'DELL *et al.* and VAN HEERDEN and JUDEK. These differences can hardly be statistical, but there is no reliable criterion for choosing between the two sets of data. It is perhaps of significance to note that the very high resolution measurements of KRISTIANSEN *et al.* [4] at $x \sim 15$ g/cm², suggest, like those of O'DELL *et al.*, that N -nuclei are appreciable less abundant than C or O nuclei.

Values of $\Gamma_{iX}(0)$ for H -nuclei have been quoted previously, WADDINGTON [8], which though of doubtful validity are still the best available. Similarly, while nuclei with $Z \geq 28$ have been reported, they appear to be so rare that little is known of their abundance except that it appears to be less or equal, but certainly not much greater, than that predicted from the cosmic abundances discussed in Section 3'4.

b) Absolute abundances. - In order to compare the relative abundances of the X -nuclei with those of α -particles and protons it is necessary to determine an absolute intensity of at least one component of the heavy nuclei. The accuracy of the determination of such an intensity is limited because:

i) The two experiments made at $x < 5$ g/cm² which have high statistical weight were both performed in emulsions exposed on a balloon whose flight duration is uncertain by as much as 20 %.

ii) The results obtained in those experiments made at $x > 5$ g/cm² have to be corrected to the top of the atmosphere by using the diffusion correction,

which, while apparently not seriously in error, must introduce a further uncertainty.

iii) The available results are from experiments performed in emulsions exposed under different conditions of solar activity, so that the intensities cannot be compared directly.

This last difficulty can be largely overcome by making the very plausible assumption that these nuclei are affected by the solar modulation process in a similar manner to the protons and α -particles. Fig. 9 shows the relationship found

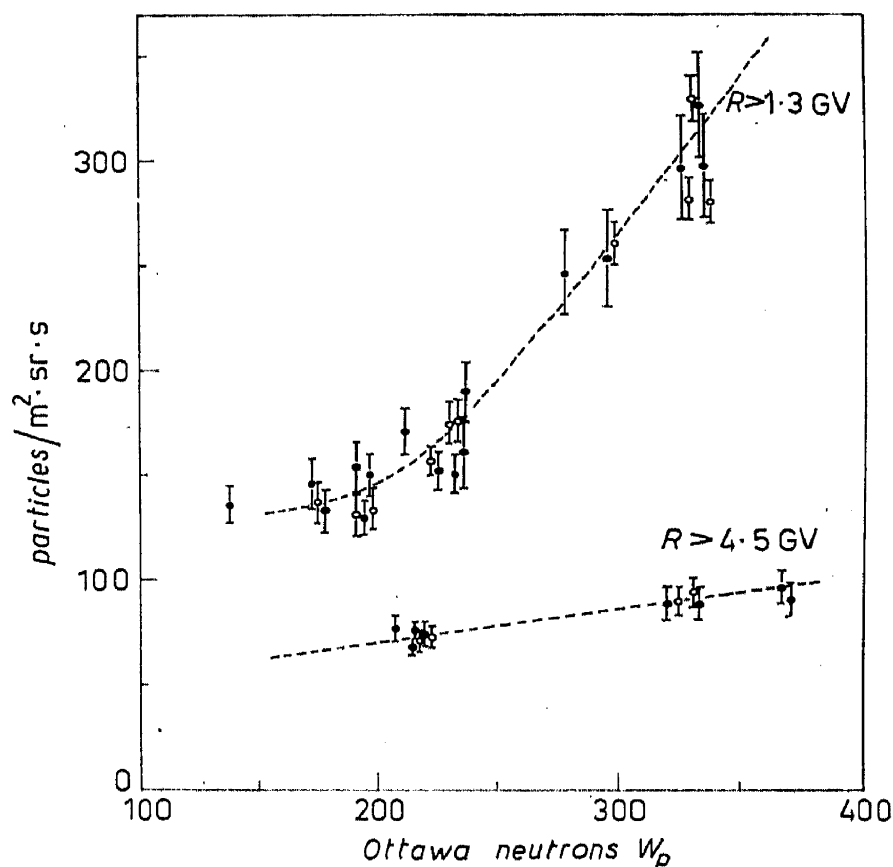


Fig. 9. - Intensities of α -particles and protons as a function of the neutron monitor value, W_p , on the date of each determination for $R > 4.5 \text{ GV}$ and $> 1.3 \text{ GV}$. Proton intensities, shown by $\circ$, are reduced by a factor of seven; $\bullet$ α -particles.

between the neutron monitor data, W_p , and the α -particle and proton intensities for $R > 4.5 \text{ GV}$ and $R > 1.3 \text{ GV}$. From these relations any intensity value can be scaled up or down to some standard W_p . If $W_p = 330$ is assumed to be this standard value, typical of solar minimum, then the corrections on each experiment are as shown in Table IV.

Table IV also gives the intensity values at the top of the atmosphere deduced in a number of experiments by using the diffusion correction, before and after correction for solar activity. Also shown are the means of the corrected values. It may be noticed that these means give $\Gamma_{LS}(0) = 0.26 \pm 0.03$ and $\Gamma_{HM}(0) =$

$= 0.38 \pm 0.04$, which are, as might be expected, in agreement with those values deduced by the diffusion approach, but not quite with those values adopted as the best values. For this reason it has been assumed that the value of J_M derived here is the most correct of the three and J_L and J_H adjusted to give these best values of $\Gamma_{LS}(0)$ and $\Gamma_{HM}(0)$.

So

$$J_L = (1.60 \pm 0.40) \quad (*) \text{ nuclei/m}^2 \cdot \text{sr} \cdot \text{s},$$

$$J_M = (5.70 \pm 0.28) \quad \text{nuclei/m}^2 \cdot \text{sr} \cdot \text{s},$$

$$J_H = (1.94 \pm 0.025) \quad (*) \text{ nuclei/m}^2 \cdot \text{sr} \cdot \text{s}.$$

TABLE IV.

Author	Observed			Date	W_p	Cor- rec- tion	Corrected		
	J_L	J_M	J_H				J_L	J_M	J_H
	nuclei/m ² sr s						nuclei/m ² sr s		
WAD- DINGTON [13]	2.30 ± 0.40	6.10 ± 0.60	2.50 ± 0.30	1954	330 ± 10	1.00	2.30 ± 0.40	6.10 ± 0.60	2.50 ± 0.30
CESTER <i>et al.</i> [19]	1.67 ± 0.39	5.52 ± 0.40	2.82 ± 0.40	not given	—	—			
ENGLER <i>et al.</i> [20]	1.68 ± 0.32	5.60 ± 0.60	2.20 ± 0.60	6.2.56	370	0.92	1.54 ± 0.29	5.15 ± 0.55	2.02 ± 0.55
GARELLI <i>et al.</i> [15]	1.87 ± 0.25	4.39 ± 0.37	1.64 ± 0.17	Sep. 58	204 ± 20	1.25 ± 0.03	2.34 ± 0.31	5.49 ± 0.46	2.05 ± 0.21
FREIER <i>et al.</i> [10]	1.90 ± 0.30	5.10 ± 0.50	1.70 ± 0.30	19.10.57	215	1.22	2.32 ± 0.37	6.22 ± 0.61	2.08 ± 0.37
						Means	2.06 ± 0.17 (*)	5.70 ± 0.28 (*)	2.17 ± 0.18 (*)

(*) Calculated from mean $\sigma/\sqrt{4}$.

Using these values, and those obtained previously for α -particles and protons, the various values of $\Gamma_{i\alpha}$ and Γ_{ip} have been calculated, Appendix I, and are tabulated in Table V.

(*) Calculated from error in J_M and $\Gamma_{MH}(0)$ or $\Gamma_{LS}(0)$.

iv) Comparison with cosmic abundances. Various attempts have been made to estimate the relative abundances of the elements in the universe. Such estimates have commonly been based on the abundances observed on the earth, in meteorites and in the atmospheres of stars, particularly the sun, SUESS and UREY [22], although recently theories of nucleogenesis have also been used to adjust these values, CAMERON [22]. It is difficult to know just how much reliance should be placed on these abundances, but it is probably reasonable to assume that, except for certain « difficult » nuclei, like neon, the

TABLE V.

	$\Gamma_{i\alpha}(R, 0)$	$\Gamma_{i\alpha}(E, 0)$	$\Gamma_{i\alpha}(U, 0)(^*)$	$\bar{A}$
p	7.0 ± 0.2	$19.8 \pm 1.2(^*)$	2.48 ± 0.15	1
L	0.018 ± 0.005	0.018 ± 0.005	0.070 ± 0.02	9.9
M	0.065 ± 0.004	0.065 ± 0.004	0.425 ± 0.026	14.0
H	0.022 ± 0.003	0.022 ± 0.003	0.46 ± 0.06	30.5

	$\Gamma_{ip}(R, 0)$	$\Gamma_{ip}(E, 0)(^*)$	$\Gamma_{ip}(U, 0)(^*)$	$\bar{A}$
α	0.143 ± 0.004	0.050 ± 0.002	0.45 ± 0.02	4
L	0.0026 ± 0.0007	0.00090 ± 0.00003	0.029 ± 0.008	9.9
M	0.0093 ± 0.0006	0.0033 ± 0.0002	0.17 ± 0.01	14.0
H	0.0031 ± 0.0004	0.0011 ± 0.0002	0.19 ± 0.03	30.5

(*) For $E \gg m_0 c^2$, see eq. (I.3).

TABLE VI. — Comparison of cosmic abundances and cosmic ray abundances.

	p	L	M	$\mathcal{H}$	$\mathcal{V}\mathcal{H}$
$\Phi_{i\alpha}$ (SUESS and UREY)	13.0	$4.7 \cdot 10^{-8}$	$1.03 \cdot 10^{-2}$	$3.6 \cdot 10^{-3}$	$2.3 \cdot 10^{-4}$
$\Phi_{i\alpha}$ (CAMERON)	6.6	$3.8 \cdot 10^{-8}$	$9.6 \cdot 10^{-3}$	$8.9 \cdot 10^{-4}$	$4.7 \cdot 10^{-5}$
$\Gamma_{i\alpha}(E)$	19.8 (*)	$1.8 \cdot 10^{-2}$	$6.5 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	$6.0 \cdot 10^{-3}$

	α	L	M	$\mathcal{H}$	$\mathcal{V}\mathcal{H}$
Φ_{ip} (SUESS and UREY)	$7.7 \cdot 10^{-2}$	$3.6 \cdot 10^{-9}$	$7.9 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$	$1.8 \cdot 10^{-5}$
Φ_{ip} (CAMERON)	$1.5 \cdot 10^{-1}$	$5.7 \cdot 10^{-9}$	$1.5 \cdot 10^{-3}$	$1.3 \cdot 10^{-4}$	$7.1 \cdot 10^{-6}$
$\Gamma_{ip}(E)(^*)$	$5.05 \cdot 10^{-2}$	$9.0 \cdot 10^{-4}$	$3.3 \cdot 10^{-3}$	$8.0 \cdot 10^{-4}$	$3.0 \cdot 10^{-4}$

(*) These values are somewhat doubtful due to the uncertain nature of the proton energy spectrum, see Sect. 3.2 and Appendix I, and the composition of the α -particles, but have been calculated assuming $E \gg m_0 c^2$.

quoted values are correct for most celestial objects to within a factor of two or three. The cosmic abundances relative to helium and hydrogen, $\Phi_{i\alpha}$ and Φ_{ip} , may be compared with the cosmic ray abundances at a given energy per nucleon, *i.e.* given velocity, $\Gamma_{i\alpha}(E)$ and $\Gamma_{ip}(E)$, and are shown in Table VI.

Here $\Gamma_{i\alpha}(E)$ is probably to be preferred, since all the values of $\Gamma_{ip}(E)$ are uncertain because of the difficulty of determining $\Gamma_{p\alpha}(E)$, and its variability when E is not much greater than m_0c^2 , Appendix I.

The most striking feature of this comparison is the considerable over-abundance of the heaviest nuclei in the cosmic radiation. This is particularly significant when it is appreciated that these nuclei are just those which would be most rapidly removed by passage through matter.

4. - The chemical distribution at low rigidities.

At rigidities which are appreciably less than 4.5 GV, nuclei with $A \approx 2Z$ no longer have relativistic velocities, and thus the ionization I becomes proportional to both Z^2 and some function of β , $f(\beta)$. The determination of charge then requires both a measurement of I and of some other parameter dependent on β . As a consequence charge resolution is generally more difficult to attain, although measurement of the energy becomes progressively more precise. With the Č-S array energy values can only be obtained when β is considerably less than unity, (Fig. 2) whereas with emulsions considerably higher energies can be measured. At rigidities less than ≈ 1.5 GV protons also cease to move with relativistic velocities, and in principle, it becomes possible to measure their energies with the Č-S array.

4.1. *Energy and rigidity spectra of helium and hydrogen nuclei.* - Alpha-particles represent the most closely studied component of the primary radiation. For this reason these nuclei will be considered in detail and then the evidence concerning the spectra of the other components examined for indications of composition changes.

The integral energy spectrum of the α -particles may be generally represented by a relation of the form

$$(5) \quad J_{\alpha^0}(>E) = \frac{\Delta}{(m_0c^2 + E)^\gamma}.$$

where E is the kinetic energy in GeV per nucleon, m_0c^2 , is the rest mass energy of a nucleon, and Δ and γ are quantities which, at least over a limited range of energy and at one particular time, are constants. Then the differential energy spectrum is given by:

$$(6) \quad \frac{dJ_{\alpha^0}}{dE} = -\gamma \frac{\Delta}{(m_0c^2 + E)^{\gamma+1}}.$$

At energies less than ≈ 0.5 GeV per nucleon the form of this spectrum has to be modified for it to represent the primary radiation. McDONALD [23] has derived an empirical function, $m(E)$, by which eq. (6) should be multiplied in order to give the spectrum observed during solar minimum. This function has the form

$$(7) \quad m(E) = 1 - \exp [-80 E^3] .$$

Fig. 10 shows the experimental results obtained during solar minimum, together with the curves represented by eq. (6) and by eq. (6) modified by $m(E)$, when $\Delta = 415$ and $\gamma = 1.5$, the values used by McDONALD.

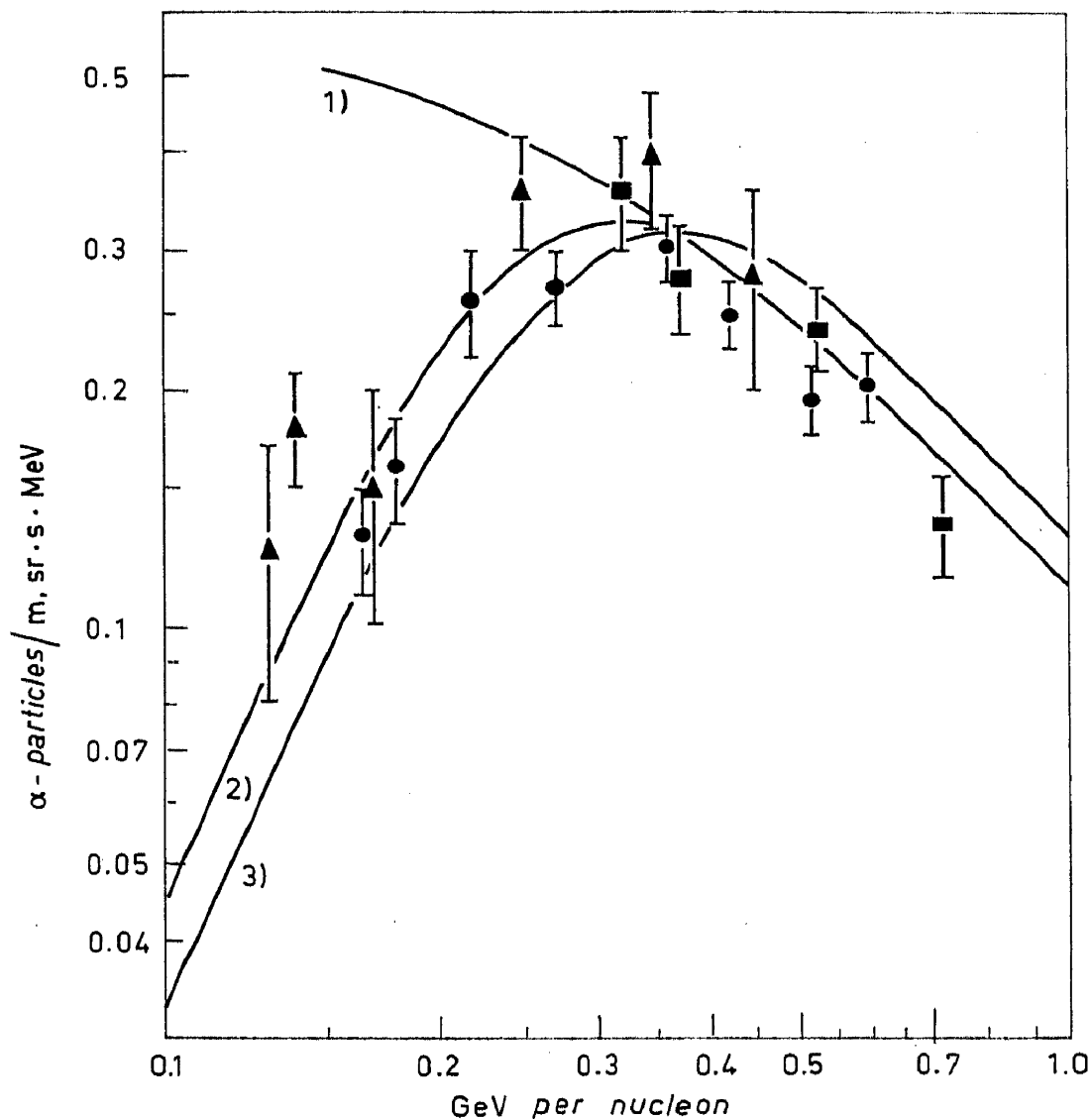


Fig. 10. — The differential energy spectra of α -particles during solar minimum. Points shown by $\blacktriangle$ are from FOWLER *et al.* [24], emulsions exposed in 1954; those by $\bullet$ and $\blacksquare$ are from McDONALD [25], Č-S array with $W_p = 335$ and 302 respectively. Curve (1) is eq. (6), curve (2) eq. (6) modified by eq. (7), and curve (3) is curve (2) adjusted for a 50% admixture of ${}^3\text{He}$ -nuclei. (See Appendix II).

The constants Δ and γ have been found from the integral and differential spectra determined from energy measurements by multiple scattering and residual range in emulsion; by pulse heights in the Č-S array; and from the comparison of intensity values at different geomagnetic threshold energies. They are found, for $0.5 < E < 5 \cdot 10^3$ GeV per nucleon to have values at solar minimum of $\Delta = 350$ and $\gamma = 1.5$, with errors of between 10 and 20 % on each. At solar maximum for $E < 7.0$ GeV per nucleon these quantities are reduced to $\Delta = 200$ and $\gamma = 1.2$. Since this reduction must be because of a greatly enhanced solar modulation effect at these times, Δ and γ presumably have the same values as at solar minimum at appreciably higher energies where these modulation effects are ineffective.

At low energies it appears that it is not the energy spectra which are similar for nuclei having different Z/A ratios, but the rigidity spectra. Equation (5) is then equivalent to an integral rigidity spectrum of the form

$$(8) \quad J_{\alpha^*}(>R) = \frac{\Delta}{[(m_0 c^2)^2 + (RZ/A)^2]^{\gamma/2}}.$$

and eq. (6) to a differential rigidity spectrum

$$(9) \quad \frac{dJ_{\alpha^*}}{dR} = -\gamma \Delta \left(\frac{Z}{A}\right)^2 \frac{R}{[(m_0 c^2)^2 + (RZ/A)^2]^{(\gamma/2)+1}}.$$

In order to represent the true spectrum observed during solar minimum eq. (9) must be multiplied by an empirical modifying function, similar to eq. (7), of the form

$$(10) \quad m(R) = 1 - \exp[-0.36 R^4].$$

This function must be applied when $R \leq 2.0$ GV.

During solar maximum a different modifying function must be used, and preferably this should be derived from some theory of the modulating mechanism, as indeed should that at solar minimum, if it is not a feature of the galactic radiation. ENGLER *et al.* [26] have suggested that the model proposed by ELLIOT [27] can adequately explain the spectra observed at solar maximum, if it is assumed that the spectrum operated on by the modulating mechanism, the galactic spectrum, is given by the differential rigidity spectrum eq. (9). This implies that the galactic spectrum observed during solar minimum is still strongly modulated, so that $m(R)$ is only of local importance. Fig. 11 shows this agreement. This model appears to be the only magnetic process so far suggested which is quantitatively able to explain the observed modulation effects. Previously it has been necessary to explain these effects in terms of

an electrostatic de-celeration mechanism, in spite of the physical implausibility of such a process, and the nearly complete absence of high-energy electrons

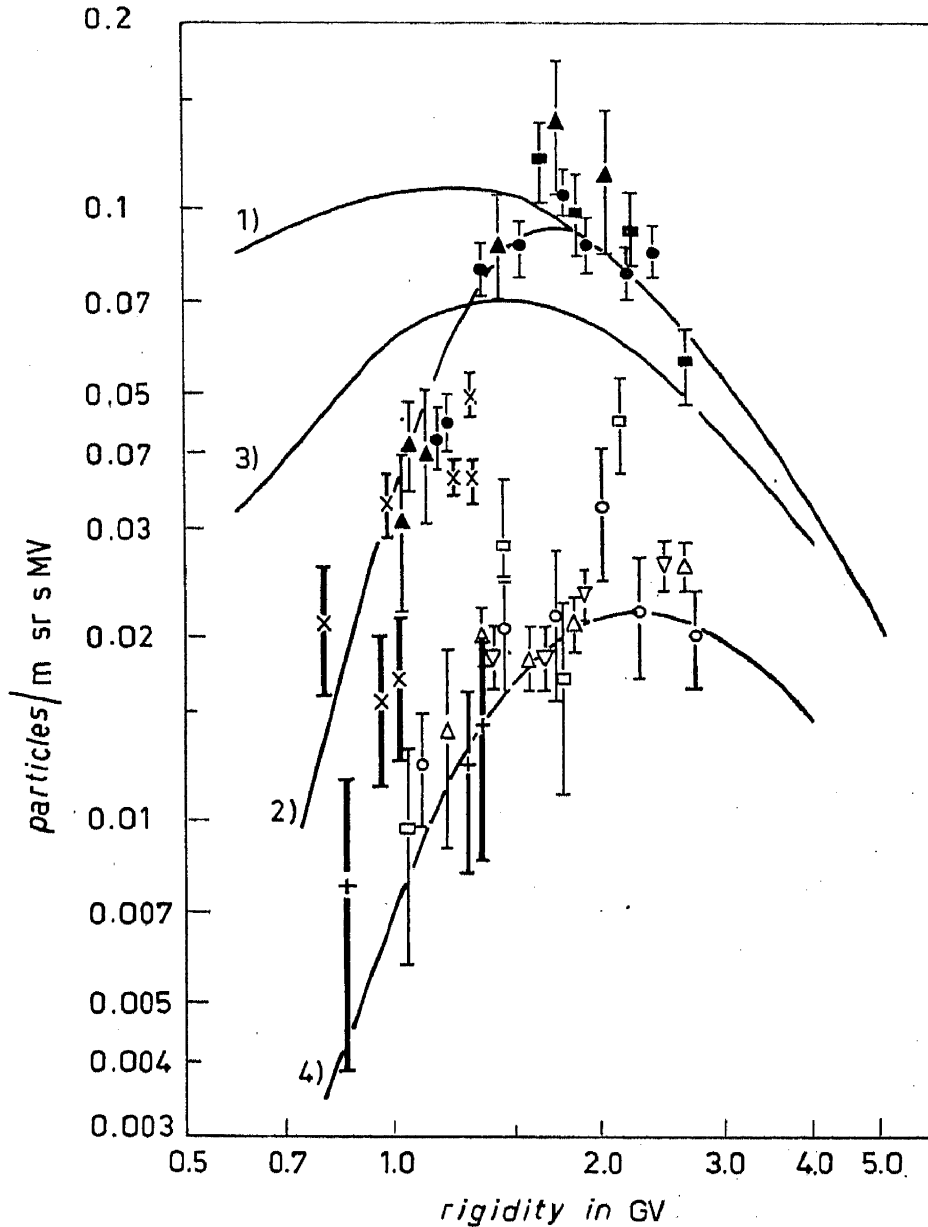


Fig. 11. - The differential rigidity spectra of α -particles and protons. Points shown solid are those on Fig. 10. Those shown by $\square$ are from AIZU *et al.* [28], emulsions with $W_p = 177$; $\circ$ are from ENGLER *et al.* [26], emulsions with $W_p = 194$; and ∇ and Δ are from McDONALD [23], Č-S array with $W_p = 191$ and 173 respectively. Points shown by $\times$ and $+$ are proton intensities divided by 7 from McDONALD and WEBBER [2,29], $\times$ at $W_p \sim 330$ and $+$ at $W_p = 237$. Curve (1) is eq. (9), curve (2) eq. (9) modified by eq. (10), curve (3) eq. (9) modified by eq. (11) using parameters appropriate to reduced solar activity and curve (4) eq. (9) modified by eq. (11) with maximum values of the parameters.

It is assumed in this magnetic model that the interplanetary medium contains a large scale solar magnetic field of predominantly dipole character generated by current systems in the solar corona, and that this field has many

small scale irregularities which act as scattering centres. The resulting modulating function should be able to explain the spectrum observed at solar minimum, as well as that at solar maximum.

ELLIOT proposes that

$$(11) \quad \left(\frac{dJ_{\alpha_0}}{dR} \right)_{\text{earth}} / \left(\frac{dJ_{\alpha_0}}{dR} \right)_{\text{galactic}} = \left(\frac{45 \cdot M_s}{r_s^2 \cdot R \cdot 10^9} \right)^r,$$

where M_s is the effective dipole moment of a coronal current system, r_s is an astronomical unit, $\text{AU} = 1.5 \cdot 10^{13} \text{ cm}$, and

$$(12) \quad Y = \frac{1}{4} [1 - \{1 + 4\alpha L [(R_0 + R)/R]^2\}^{1/2}],$$

where αL and R_0 are parameters which respectively relate to an absorption probability and the scale and strength of magnetic inhomogeneities in the solar system.

ELLIOT suggests values of the various parameters at a time of reduced, although not negligible, solar activity, and these have been used to calculate $(dJ_{\alpha_0}/dR)_{\text{earth}}$ assuming that eq. (9) gives $(dJ_{\alpha_0}/dR)_{\text{galactic}}$. Figure 11 shows the lack of agreement between this calculated spectrum (3), and that observed experimentally (2), and suggests that it will be difficult to find a function to modulate curve (1) to curve (2). This may be due either to an inadequacy in the assumed parameters or the model, or to the failure of eq. (9) to represent the true galactic spectrum. It is possible that the parameters given by ELLIOT can be adjusted sufficiently so that the *modified* solar minimum spectrum, curve (2), could be modulated to give the solar maximum spectra curve (4). This implies that the galactic spectrum is represented by eq. (9) modified by $m(E)$. The essential point to note is that the true galactic spectrum in this range of rigidities must, therefore, be regarded as being somewhat uncertain, since it is not clear whether the modulating functions observed at solar minimum are characteristic of the galactic radiation or are imposed near the solar system.

An additional difficulty in the interpretation of the experimental spectra is the recent demonstration by APPA RAO [30] of an appreciable abundance of ^3He among the α -particles. As a consequence experimental measurements of energy are in error, and threshold values are also ambiguous. The corrections introduced in the energy spectrum are discussed in Appendix II, where they are shown to be rather small. Furthermore, the error made in calculating energies is nearly cancelled by then calculating the rigidity, so a rigidity spectrum appears to be less subject to errors from this cause. Equations (8) and (9) do however involve Z/A , and thus the deduced constants are still erroneous to a small extent.

Figure 11 also shows the experimental intensity values obtained by McDONALD and WEBBER [2] for slow protons, normalized to the α -particles by dividing by the value of $I_{\alpha}(R)$ found previously at higher rigidities. These results certainly indicate that the protons have a similar rigidity spectrum to the α -particles in this range. However, they have not been checked in any way by other experiments, and should be accepted with some caution till such a confirmation is available. Experiments by NEHER [31] and by WINCKLER and ANDERSON [32] during solar minimum, and by VOGT [33] during 1960, indicate that at still lower rigidities, < 0.7 GV, the proton intensity must commence to rise again, though what the α -particles do in this rigidity range is not known. These results, if confirmed, may suggest a steady solar component of very low-energy particles.

4.2. *Low-rigidity $Z \geq 3$ nuclei.* — The situation with regard to the rigidity spectra of these nuclei is obscure, being bedevilled by experimental inconsistencies. EVANS [34] has obtained an energy spectrum for all $Z \geq 3$ nuclei in a stack exposed at $W_p = 335$, which is fitted by eq. (9) modified by $m(R)$ with $\Delta = 33.4$. TAMAI [35], in the *same* stack, finds evidence for peaks in the differential energy spectra which depend on Z ; more L -nuclei than M -nuclei and an α -particle spectrum which resemble that observed at solar maximum. AIZU *et al.* [28] find at solar maximum a total intensity which is, as would be expected, about half that observed by EVANS, but differential intensities at energies less than ~ 700 MeV per nucleon which are approximately equal in magnitude. Unlike TAMAI they find L -nuclei are about half as abundant as M -nuclei.

Figure 12 shows the rigidity spectra found by EVANS [34] for $Z \geq 3$ nuclei, and those found by AIZU *et al.* [28] for L and M -nuclei. It can be seen that these spectra are similar in form to those spectra represented by eq. (9) modified by eq. (10) or eq. (11) as appropriate. However, AIZU *et al.* point out that an analysis of their results suggests, with rather doubtful statistical significance, that I_{LM} varies in this range of rigidities, a conclusion which is not supported by EVANS, who does not observe any broadening of the peak in the differential spectrum. It has been shown by various authors, *e.g.* TERASHIMA [36] that if the true galactic spectrum is that observed at solar minimum, then due to ionization energy loss experienced by the nuclei as they traverse the interstellar medium, the shapes of the differential rigidity spectra of the various components will be different, with the peak in the spectrum moving to higher R as Z increases. This assumes that all nuclei pass through the same amount of matter from the source to the earth, which may well not be the case, see Section 8. Thus the definite observation of such an effect would be of considerable importance since it would be both an indication of the form of the galactic spectrum and of the mode of propagation in interstellar space.

The value of Γ_{LS} in this range of rigidities is rather uncertain. FICHTEL [37] finds, for those particles with $R > 1.8$ GV, $\Gamma_{LS} = 0.35 \pm 0.06$ under 8.8 g/cm^2 ,

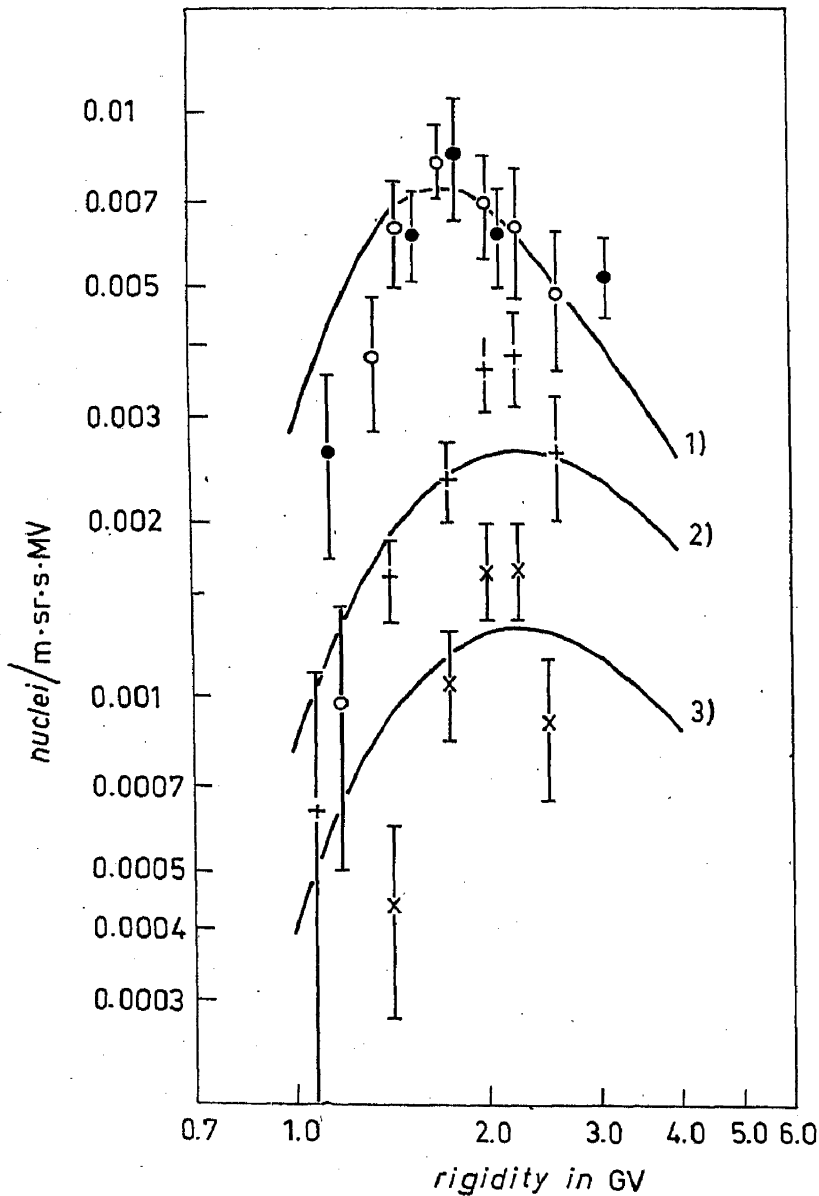


Fig. 12. - The differential rigidity spectrum of $Z \geq 3$ nuclei. Points shown as $\circ$ and $\bullet$ are for all nuclei with $Z \leq 3$ from EVANS [34] at $W_p = 335$. Those points shown by $+$ are for M -nuclei, and those by $\times$ for L -nuclei, from AIZU *et al.* [28] at $W_p = 177$. Curve (1) is eq. (9) modified by eq. (10) with $\Delta = 33.4$, while curves (2) and (3) are for eq. (9) modified by eq. (11) with maximum values of the parameters, and Δ determined from the total intensity.

which is similar to the value found for $R > 4.5$ GV. However, about 50% of Fichtel's nuclei have $R > 4.5$ GV and so this value of Γ_{LS} is not very sensitive to variations at low rigidities. AIZU *et al.* [28] find, on the basis of an extrapolation through 7.6 g/cm^2 , $\Gamma_{LS}(0) = 0.33 \pm 0.04$ at the top of the atmosphere for nuclei with $1.3 < R < 2.6$ GV. This is probably the best available evidence that Γ_{LS} does not vary greatly at these low rigidities from the value observed at higher rigidities although it is significantly higher. TAMAI [35], on the other hand, finds $\Gamma_{LS} > 1.0$ under 7.0 g/cm^2 for $R < 2.5$ GV, in sharp disagreement with the results given above. These are the only published measurements made under less than 12 g/cm^2 , but there are a number of experiments performed deeper in the atmosphere which report results similar to those of TAMAI [35]. Experiments in detectors exposed under small amounts of residual atmosphere will be needed to resolve this controversy.

4'3. *Special components observed at low rigidities.* - At these low rigidities there is experimental evidence for two other

components of the primary radiation which have not, as yet, been observed at higher rigidities.

a) *Primary electrons* (*). – Until recently the only evidence regarding the possible existence of primary electrons was the negative result of CRITCHFIELD *et al.* [38], who quoted an upper limit to the intensity of electrons with $E > 1$ GeV of 13 electrons/m²·sr·s. Recently, however, MEYER and VOGT [39] have used a counter telescope with absorber to give finite limits on a definite electron intensity. These authors claim that for $25 < E < 100$ MeV the intensity of vertically incident electrons is $(280 \div 310)$ electrons/m²·sr·s, for $100 < E < 1300$ MeV it is $(35 \div 110)$ electrons/m²·sr·s; while for $E > 1.3$ GeV it is $(0 \div 80)$ electrons/m²·sr·s. In addition EARL [40] has used a cloud chamber technique to show that $N_e/N_p = (3 \pm 1)\%$ for $E > 0.5$ GeV. These values may be compared with the upper limit of 0.1% derived by DUTHIE *et al.* [41] for the ratio of primary γ -rays and/or electrons to protons with $E > 470$ GeV. These finite intensities of low energy electrons compare well with those that would be expected from the observations of radio intensities in the galaxy. Following GINZBURG [42] Fig. 13 shows the density of electrons with $E > 1$ GeV plotted as a function of the near galactic magnetic field for various values of α , the exponent of the radio frequency spectrum. It can be seen that for $H \approx 5 \cdot 10^{-6}$ gauss the density of electrons should be a few per cent of that of the cosmic radiation. However, it may be noted that these experimental determinations were made at a time of solar maximum, and it would therefore be more realistic to compare densities at a given magnetic rigidity rather than energy. Because the rigidity of an electron at a given energy is lower than that of a proton these ratios and intensities are presumably lower limits to the galactic values.

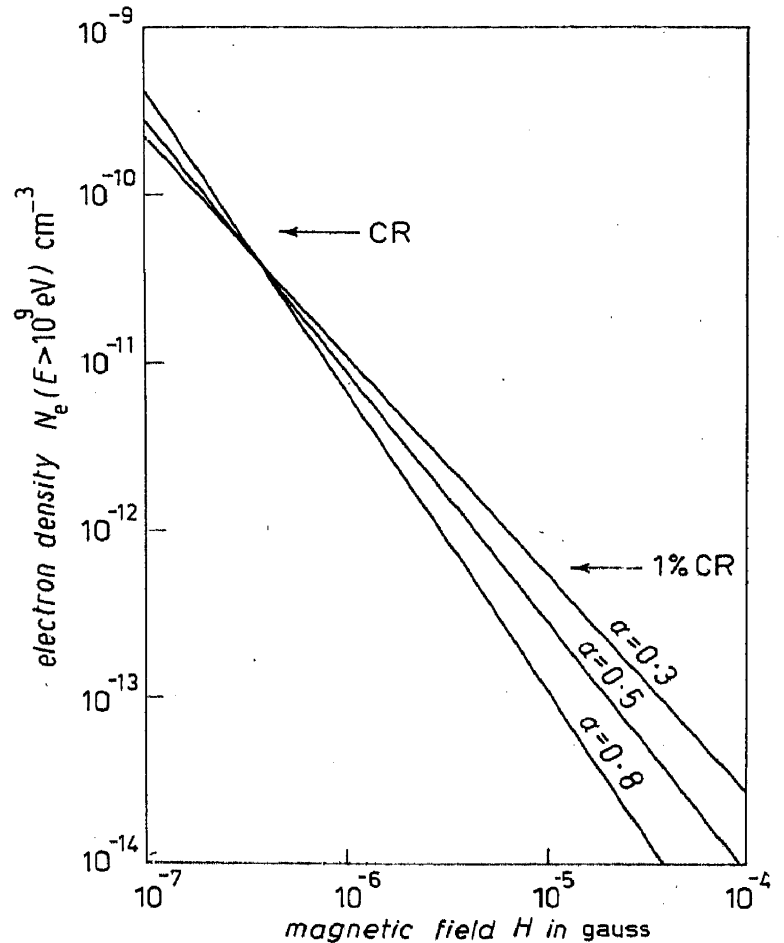


Fig. 13. – The density of electrons with an energy above 1 GeV as a function of the magnetic field for different values of α , the exponent of the radio frequency intensity relation, $I = \text{const } \nu^{-\alpha}$. Calculated from the data given by GINZBURG [42].

(*) See DAYTON in these proceedings.

b) ^3He -nuclei. — Measurements in emulsions of range and multiple scattering on stopping α -particles have enabled APPA RAO [30] to distinguish between ^3He and ^4He nuclei. For $0.2 < E < 0.4$ GeV per nucleon he finds an abundance ratio $^3\text{He}/^4\text{He} = 0.70 \pm 0.16$, while for $1.3 < R < 1.6$ BV a ratio of 0.57 ± 0.18 . As would be expected these values are not significantly different. The determination of this ratio is of particular interest because of the low cosmic abundance of ^3He nuclei. The significance of this will be discussed in Section 8.

5. — The chemical distribution at high rigidities.

At rigidities appreciably above 4.5 GV the data on the relative abundances of the various components become rather sparse, mainly because of the difficulty of obtaining results of significant statistical weight. At > 16 GV, the vertical threshold rigidity near the equator, measurements of the intensities of protons, α -particles, M and H -nuclei are all consistent with the relative abundances being similar to those at 4.5 GV, WADDINGTON [8], *i.e.* with having an energy or rigidity spectrum whose power index, γ , is 1.5. Notice from eqs. (5) and (8) that when R or E is much greater than m_0c^2 both $J(E)$ and $J(R)$ obey a power law with the same index γ .

Systematic observations on the intensities of high-rigidity protons and α -particles have been extended up to $\approx 5 \cdot 10^3$ GV from studies of high-energy nuclear interactions, « jets » observed in emulsions, Plate II. Estimates of the energy, and thus the rigidity, of these particles depend on the characteristics of the individual interactions, and for any one event may be incorrect by a factor of two or more. Consequently any detailed structure in the spectrum will be obscured and at best can only appear as a gradual trend.

Unless a jet is produced by a particle having an energy appreciably greater than 10^3 GeV per nucleon it is unlikely to be detected by its secondary effects but the interaction itself must be located. It is thus necessary to detect a point rather than a line in the emulsion and consequently, it is more difficult to obtain a sample of appreciable statistical weight of these « low »-energy jets than it is for jets produced by the still rarer particles of higher energy. There is thus a gap, between 16 GV and at least 10^3 GV, in which very few measurements have been made, simply due to this difficulty of detecting the events.

Figure 14 shows the intensities obtained by various workers of particles with rigidities above some given limit. The energies of the individual events were estimated either from an estimate of the energy in the electromagnetic cascade, or from the geometrical configuration of the nuclear particles in the jet. These results are consistent with the derived integral rigidity spectrum, eq. (8), with $\gamma = 1.5$; and show that $\Gamma_{\text{px}}(R)$ has the same value as at lower rigidities. In addition, ENGLER and HABER-SCHAIM [43] found, for proton-

induced events, $\gamma = 1.4 \pm 0.2$ between 65 and 500 GV. Thus up to $5 \cdot 10^3$ GV there is no evidence for any change in γ from the value observed at lower rigidities.

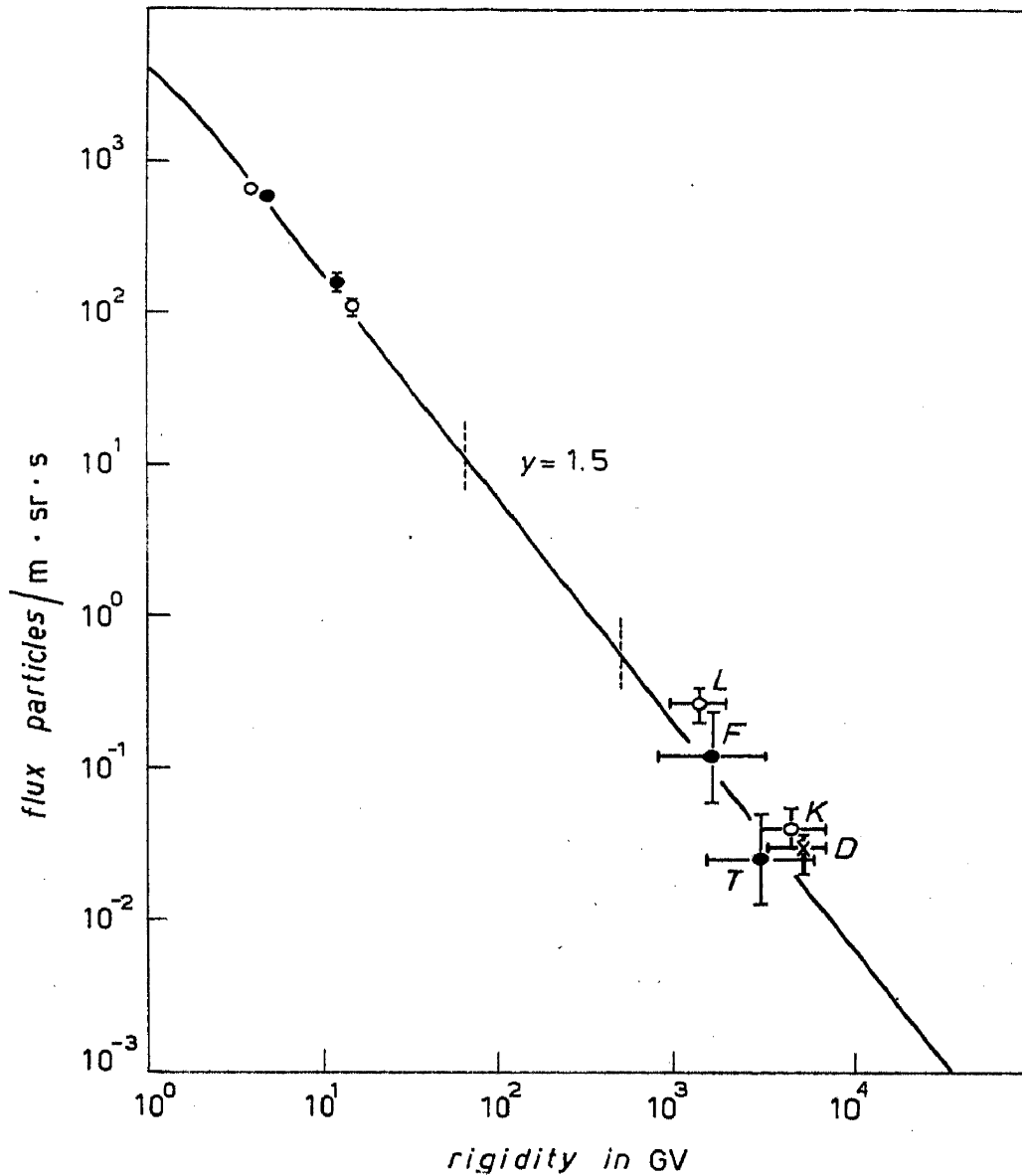


Fig. 14. — The integral rigidity spectrum of protons (O) and α -particles ($\times 7$) (●) together with eq. (8) when $\gamma = 1.5$. Between the rigidity values shown by vertical dotted lines ENGLER and HABER-SCHAIM [43] determined γ experimentally. Points shown for proton intensities are *L*, LAL [44]; *K*, KAPLON *et al.* [45] and *D*, DUTHIE *et al.* [41], while for α -particle intensities *F*, FOWLER and WADDINGTON [46] and *T*, LOHRMANN and TEUCHER [47].

Individual events having rigidities of up to 10^5 GV have been observed in emulsions, but not in a systematic manner because of the extremely low intensities. The interpretation of such events is particularly difficult due to the large fluctuations in the apparent energy that can occur for individual events. As a consequence of these fluctuations and of the steepness of the rigidity spectrum, events having particular characteristics which indicate a certain primary rigidity, can instead be the result of a fluctuation of some of the more

abundant lower rigidity particles. Similar considerations apply to extensive air showers.

For $Z \geq 3$ nuclei, JAIN *et al.* [48] find that, up to 200 GV (=100 GeV per nucleon), the rigidity spectra of the M and H -nuclei are still represented by $\gamma = 1.5$. No systematic observations of still higher rigidity heavy nuclei have been made, although individual nuclei having rigidities of up to 10^5 GV have been observed producing jets in emulsions. As in the case of the very-high energy proton and α -particle-induced jets, the number of reported events



Plate II. - Two proton-induced jets observed in emulsion, one of ≈ 200 GeV and the other of ≈ 1200 GeV. Notice the different geometrical configuration of the tracks of the fast particles in the two cases, and that in both the particles spread out so rapidly that these events can only be detected near to the star.

is not inconsistent with that to be expected if $\gamma = 1.5$ at these rigidities, but does not define γ with any precision.

Above 10^5 GV there is no direct information on the relative abundances of the various nuclei. However, it should perhaps be stressed that there is *no* convincing experimental evidence to prove that the relative abundances vary appreciably with rigidity, nor, apart from the variation in $I_{p\alpha}(E)$ at lower energies, with energy, over the whole range in which observations have been made.

6. - The energy distribution at very high rigidities.

At energies above 10^5 GeV per nucleon information on the energy spectrum can only be obtained from observations made deep in the atmosphere on the secondary products of the very energetic primary particles. Apart from those experiments in emulsions exposed on aeroplanes, these measurements are sensitive only to the total kinetic energy of the particle (*), U , and thus if the chemical composition is the same as at lower energies, nearly half the events at any given apparent energy must be produced by $Z \geq 2$ nuclei, Table V.

Measurements have been made by ALY *et al.* [49] on nuclear induced jets produced in emulsions interleaved with sheets of high Z material exposed both at aeroplane altitudes, ≈ 220 g/cm² and at balloon altitudes, ≈ 10 g/cm². These events were ones in which more than 500 GeV were originally given to the π^0 -mesons created in the nuclear interactions, and were detected from the resulting electromagnetic cascade, whose development was enhanced by the sheets of high- Z material. Assuming charge independence and an elasticity in the interaction of between 50 and 75 %, the energies of the primary particles should have averaged some 6 to 12 times that in the cascade. The resultant energy spectrum, which is of the form $J(U) \propto U^{-\gamma}$, has $\gamma = 2.0 \pm 0.2$. This is probably, but not necessarily, the spectrum of the primary radiation, since equating the two assumes, for example, that the fraction of the primary energy that goes into π^0 -mesons is constant over the energy range. This spectrum may be normalized to a primary spectrum by using the intensity of $(3 \pm 1) \cdot 10^{-2}$ protons capable of producing electromagnetic cascades containing more than 500 GeV energy per m²·sr·s found by DUTHIE *et al.* [41] by an

(*) This statement requires some justification since it is not precisely valid. At these energies a nucleus may be regarded as consisting of A unrelated nucleons, each of which is capable of producing an air shower of size $n \propto E^\beta$, where E is the energy of a nucleon, and β is some constant at sea level. Hence the shower observed will have a size $N \propto A E^\beta$, and since $U = A E$, $N \propto A^{1-\beta} U^\beta$.

Now from eq. (15), $U \propto N^{0.85}$ so β is approximately unity and thus $N \propto U$. This argument will not be justified for showers appreciably above sea level, since N is far from linear with energy under these conditions.

a) μ -mesons. BARRETT *et al.* [50] have used measurements of the intensity of single μ -mesons detected under 1600 m w.e. and of multiple pairs of μ -mesons under the same thickness of matter, to derive two points on the integral energy spectrum of the primary particles. The derivation of these intensity and energy values depends on relatively plausible but not rigorously justifiable assumptions, and they must therefore be regarded with some reservations. These values, which are shown in Fig. 15, are presumably not seriously affected by the discrepancies in the reported intensities of higher-energy μ -mesons recently revealed by DUTHIE *et al.* [41] from observations of γ -rays in emulsions and by BARTON [56] from some re-measurements of these values.

b) Extensive air showers. These are observed with arrays of counters which measure directly a small sample of the particles, predominately electrons and μ -mesons, produced in the cascade process. These measurements are used to estimate the total number of particles present in the cascade when it strikes the array, and this number is then related to the energy of the producing particles by assuming some specific model for the physical process of transferring energy to the cascade. As for the case of the underground μ -mesons these calculations can hardly be rigorously justified but appear plausible.

At energies of $\approx 10^6$ GeV a point on the integral energy spectrum can be obtained by integrating the energy lost by air showers of known total energy over the whole atmosphere. Such points have been calculated by GREISEN [52] and COCCONI [53].

The energy spectrum has been calculated from the observed number spectrum by making the simple assumption, which almost certainly leads to an overestimate of the energy, and thus a lower limit of γ , that the average energy per particle in a shower is independent of the size and equal to that found for the air showers of known total energy, these being the smallest considered. This average energy is calculated by COCCONI [53] to be 12 GeV, and assuming that the integral number spectrum is given by

$$(13) \quad F(>N) = 0.59 N^{-(0.815+0.060 \log N)} \text{ showers/m}^2 \cdot \text{sr} \cdot \text{s},$$

he finds an integral energy spectrum, which he presumes to be an upper bound to the true spectrum, and is shown in Fig. 15. GREISEN presents rather similar data from which an energy spectrum can be calculated which takes into account the variation of the average energy with size. He derives the following integral number spectrum

$$(14) \quad F(>N) = (3 \pm 0.6) \cdot 10^{-8} (N/10^6)^{-(1.66+0.06 \log (N/10^6))} \text{ showers/m}^2 \cdot \text{sr} \cdot \text{s},$$

which is claimed to be valid from $N=10^3$ particles up to the largest showers yet recorded, $>10^9$. Also it is concluded that in a shower with $N=5 \cdot 10^5$ particles at sea level the average primary energy of such showers is (15 ± 3) GeV

per particle at sea level. Finally observations by CHUDAKOV *et al.* [57] are quoted to show that the average energy per particle is greatest in small showers, and varies as $N^{0.15}$, so that

$$(15) \quad U = N \cdot 15 \cdot \left(\frac{5 \cdot 10^5}{N} \right)^{0.15} \text{ GeV}.$$

The resulting spectrum is shown in Fig. 15, where the uncertainties are solely those due to the ± 0.6 term.

Perhaps the most extensive experimental data in this energy range are that of the MIT group, CLARK *et al.* [55]. These authors assume a specific model for shower development, which, unfortunately, is not available in the literature, and, by taking into account fluctuations in the depth at which the first interactions occur, find an integral energy spectrum of the primary nuclei, assumed to be all protons, of

$$(16) \quad J(> U) = (8.1 \pm 3.1) \cdot 10^6 U^{-2.17} \text{ protons/m}^2 \cdot \text{sr} \cdot \text{s},$$

when $3 \cdot 10^6 < U < 10^9$ GeV. Although this spectrum is somewhat depressed relative to that derived by GREISEN [54] this may be due to the allowance made for the effect of fluctuations. Obviously if many of the primary particles producing these events are heavy nuclei these fluctuations will be greatly reduced, and the spectrum brought into closer agreement with Greisen's.

The largest showers yet observed contained considerably more than 10^9 particles at sea level, which corresponds to a primary proton energy near to 10^{10} GeV. Furthermore, since at least two of these extremely large events have been observed with zenith angles of nearly 60° , if they were initiated by heavy primaries the energy estimates have to be greatly raised because then they could hardly be the result of a fluctuation. Consequently it appears that there must be primary particles with rigidities $> 10^9$ GV, regardless of whether they are protons or heavier nuclei. There is no evidence for the existence of an upper limit to the energy spectrum, since still larger showers should be so infrequent that they would not yet have been detected by present E.A.S. arrays.

All this data are summarized in Fig. 15, from which it appears that there are serious discrepancies between the various spectra, which, at some energies, amount to as much as a factor of 10^2 in the intensity. Nevertheless over the range of energies between 10^3 and 10^9 GeV it is possible to assume that γ lies between 2.0 and 2.2, although it is impossible to verify whether there are any serious discontinuities or abrupt changes of slope. PETERS [58] has suggested, from an examination of air shower size data obtained at mountain altitudes, that there is a discontinuity for showers with more than $\sim 4 \cdot 10^5$ particles, which, at that altitude, would correspond to a primary energy of $\sim 10^6$ GeV. Such a discontinuity could be due to a change in the primary spectrum, or

to a change in the physical processes involved in these very-high energy nuclear and electromagnetic interactions. A more detailed investigation of this point is undoubtedly required. Equally the apparent consistency of γ down to $5 \cdot 10^3$ GeV shown in Fig. 15 might be taken as indicating that γ changes rather abruptly from 1.5 to 2.0 at this energy, but the spread of the results is far too great to justify such a conclusion. It would appear that the energy spectrum should also be investigated with more care in this region, but, as pointed out previously, the experimental difficulties are particularly severe at this energy. At the present time it would seem that the majority, if not all, of the experimental evidence suggests that γ varies slowly and smoothly from 1.5 to about 2.2 between 10 GeV and 10^9 GeV.

7. - The spatial distribution.

Any anisotropy in the directions of motion of galactic particles arriving in the vicinity of the earth should produce a sidereal-diurnal variation in the intensity recorded at the earth. It appears that if such variations exist they are generally small, less than 2 or 3 %, and as a result investigations of these effects have been made almost exclusively with counting equipment placed deep in the atmosphere. Such equipment, *e.g.*, neutron monitors, meson counters, E.A.S. arrays, can only observe the secondary products of those primary particles which have an energy above some limiting value. Consequently an observed constancy in the intensity may not reflect a similar constancy in the charge distribution, but only in the total energy input. For example, at low energies a neutron monitor, because of its small efficiency of detecting any one primary particle, would observe little difference between the arrival of one iron nucleus or of its 56 component nucleons, while at high energies an air shower array, which is only sensitive to the total energy, cannot distinguish between an iron nucleus and a nucleon of the same total energy.

Experimentally the detection of a definite sidereal variation in the intensity has proved extremely difficult, and it still appears doubtful whether any true sidereal variation occurs. DORMAN, in 1959 [59], summarized the situation by stating that if sidereal variations do exist, their amplitude is ≤ 0.02 % as measured by meson counters, and ≤ 1 % as measured by air shower arrays. It is apparent, therefore, that the energy content, and probably the particle density, is very nearly isotropic in space for energies between 5 and 10^7 GeV. CLARK *et al.* [55] state that they have no evidence for any obvious anisotropy even at the highest energies, 10^9 GeV. At energies less than a few GeV any sidereal effects are obscured by solar-diurnal and irregular variations of >1 %, as observed by neutron monitors, and are anyway intrinsically improbable due to the dominant nature of local magnetic fields on particles of such low rigidities.

This remarkably high degree of isotropy, applying even to particles of such great rigidity that their radii of curvature in possible galactic magnetic fields appear to approach the dimensions of the galaxy, must be explained by any acceptable theory of the origin and propagation of these particles. Furthermore, it should be noted that the limits placed on any possible sidereal variation of the majority of the radiations are at least an order of magnitude less than the expected Compton-Getting effect, which would arise as a result of the rotation of the galaxy in inter-galactic space, DORMAN [60]. Thus these particles must partake of the mass motion in the galaxy, which is strong evidence for the presence of appreciable magnetic fields which « bind » them to the galaxy.

8. – Significance of the results to theories of the origin.

The value of the experimental observations summarized thus far lies in the restrictions that they place on the possible theories of the origin and propagation of the particles in the cosmic radiation. These theories may be broadly gathered into three groups:

- i) The particles originate in the sun and are confined to the solar system.
- ii) The particles originate in the galaxy, and are confined within it.
- iii) The particles originate in, or between, the galaxies but are not confined to them.

Almost certainly particles whose origin could be classified in each of these groups are received at the earth, but equally certainly, their relative importance, or intensities, are greatly different.

8.1. *Solar origin.* – Particles, including nuclei with $Z \geq 2$, definitely originate in or near the sun during solar disturbances, although apparently with relative abundances which differ appreciably from those usually observed in the primary radiation, $I_{Mp}(R)$ being only 10% of that observed in the cosmic radiation (FICHTEL and GUSS [61]). Exceptionally these particles may have rigidities as high as 50 GV, but more usual values are of the order of $(0.1 \div 1.0)$ GV, with, a rigidity spectrum typically proportional to $R^{-(4 \div 6)}$. Whether any appreciable number originate at times other than during these solar disturbances is uncertain, though the results reported by VOGT [33] suggest a steady flux of very low rigidity (< 0.7 GV) protons. Some of the arguments against solar particles making up any considerable part of the radiation above 1 GV are:

- a) The absence of any appreciable solar diurnal variations. For this to be consistent with a solar origin there must be a permanent and efficient mechanism to render the particles isotropic, similar to that which acts on the

high energy solar flare particles after the first few minutes of a flare increase. However, particles with energies less than one or two GeV per nucleon are not observed at sea level, so that they could exhibit considerable diurnal variations which have not yet been detected.

b) The overabundance of the high Z elements. This could be due to a preferential acceleration mechanism, but is not consistent with the observed paucity of these elements in solar flare radiation.

c) The failure to detect any transition in either the chemical or the rigidity distributions from solar to galactic particles, although particles with rigidities much greater than 10^2 GV, which corresponds to a particle having $r=1$ AU in a field of $2 \cdot 10^{-5}$ gauss, can hardly be of solar origin, and still be isotropic. Unfortunately, this rigidity limit already lies beyond the range in which these distributions are relatively well known.

d) FAN *et al.* [62] failed to find any appreciable increase in particle density over 0.1 AU on the Venus space probe. These results were only preliminary, however, and in any case an increase might only be expected considerably closer to the source, if at all.

None of these arguments are conclusive, therefore. Hence the possibility that solar particles do make up an appreciable fraction of at least the low rigidity, and thus of all the particles, should not be entirely neglected, in spite of the greater intellectual appeal of the galactic theories, which relate the problem of the origin to much other astrophysical evidence.

8'2. *Galactic origin.* — Certainly some, and probably the majority, of those particles commonly called cosmic rays, are of galactic origin, and must therefore have been propagated through the interstellar medium.

Neither the places of origin, nor the mechanisms of acceleration, are understood with any certainty. However, it appears fairly convincing that large numbers of energetic particles are to be found in those galactic objects which are strong sources of non-thermal electromagnetic radiation, since this radiation is generally considered to be produced by the motion of energetic electrons moving in magnetic fields, synchrotron radiation. Such sources are the central nucleus of the galaxy and supernovae remnants like the *Crab nebula* and *Cassiopeia A*. Whether these supernovae remnants represent the most important sources of cosmic ray particles is not known, although the overabundance of heavy nuclei observed in the cosmic radiation is often quoted as supporting this view, since the nucleogenesis of heavy elements is believed to be important in the causation of supernovae.

Similarly the mode of propagation of the particles produced in these, or other, sources, is not well known. The polarization of star light, the existence of a radio emission background in the galaxy, and the absence of a measurable

Compton-Getting effect, strongly indicates the presence of extensive magnetic fields in the galaxy. Attempts to measure the strengths of these fields directly have so far only succeeded in setting an upper limit on the permissible value. Little appears to be known about the configuration of these fields, although the polarization evidence indicates a degree of large scale regularity along the galactic arms. On the other hand interstellar matter appears to be very irregularly distributed, forming « clouds » of gas, and if magnetic fields are associated with these clouds, many, comparatively small-scale irregularities must be expected. For this reason the propagation of the cosmic ray particles is often regarded as a diffusion process, with the gas clouds, and their associated magnetic fields, acting as scattering centres. Such a mode of propagation, possibly, but not necessarily, combined with a degree of confinement in the galaxy, would explain an isotropy of the cosmic radiation in spite of a non-isotropic distribution of sources. It is possible that particles are accelerated by the Fermi mechanism, while travelling in interstellar space, but the astronomical evidence suggests that the scale and relative velocities of the gas clouds are insufficient for the efficient operation of such a mechanism. On the other hand, it is difficult to visualize any mechanism whose dimensions are restricted to those of a nebula, being able to accelerate particles up to 10^{10} GeV.

If the mode of propagation is a diffusion without acceleration, from a relatively concentrated region of sources, then the equations for the concentration, N_i , of nuclei of type i , are:

$$(17) \quad \frac{\partial N_i}{\partial t} - \text{div}(D \nabla N_i) + \frac{N_i}{T_i} = Q_i + \sum_{j < i} P_{ij} \frac{N_j}{T_j},$$

where D is the diffusion coefficient, $= lv/3$ with l the effective distance between magnetic inhomogeneities and v the particle velocity, $\sim c$; T_i is the absorption lifetime, *i.e.* the time a particle takes to travel an absorption mean free path λ_i ; Q_i is the source density of i type nuclei, and P_{ij} is the fragmentation parameter, *i.e.* the average number of j type nuclei produced by the *absorption* of an i type nucleus. Solutions of these equations, and of those applicable to the simpler one-dimensional case, are given by GINZBURG and SYROVAT'SKII [63] and are of the form

$$(18) \quad N_i = f(t_i, Q_i).$$

Thus, in principle, if any value of Q_i is known, t_i , the mean lifetime of i type nuclei, can be found, and thus, since t_i is related to t_j , all other values of Q_j . In practice, T_i , P_{ij} and D have to be known, and there are appreciable experimental uncertainties regarding all these parameters. In the one-dimensional case $D = 0$ and the analysis is usually made in terms of s , the distance traversed in g/cm^2 , $= \sigma ct$, where σ is the density of the medium, and $\lambda_i = \sigma c T_i$.

In this case each particle has passed through the *same* value of s , *i.e.* the same t , whereas in the three-dimensional case t is only an *average* value, and different particles may have traversed different amounts of matter, permitting the concept that certain particles may be «younger» than others. Thus for example iron nuclei may, in general, be younger than protons.

Usually Q_z is assumed to be zero, and the equations are solved for each component in turn. Due to the uncertainties in N_i , λ_i and P_{ij} such solutions are often in considerable disagreement. For example HAYAKAWA *et al.* [7], who assume $\Gamma_{Ls} = 0.23$ find in the one dimensional case $s = 3 \text{ g/cm}^2$, while GINZBURG and SYROVAT'SKII [63], who assume a similar value of Γ_{Ls} , find, also in the one-dimensional case, $s = 6.5 \text{ g/cm}^2$. These values may be compared with that found by APPA RAO [30] of $\sim 12 \text{ g/cm}^2$ by assuming $Q_{He} = 0$. This appreciably larger value could be due to errors in the parameters, or may reflect the greater «age» of these lighter nuclei. This latter explanation is also suggested by the proposal made by HOYLE [64] to account for the high energy electrons that produce synchrotron radiation, that cosmic ray protons have passed through some 30 g/cm^2 of matter. Further, while HAYAKAWA *et al.* [7] state that the available data are quite inconsistent with all $Q_i = 0$ apart from $Q_{\gamma, H}$, GINZBURG and SYROVAT'SKII find this an acceptable hypothesis.

Thus these parameters need to be experimentally determined with enhanced accuracy and other components studied for which $Q_i = 0$. Some such components are ^2H , ^3He , all the L -nuclei, preferably with each element considered separately, and F -nuclei. In addition it would be of great interest to detect examples of radio-active nuclei, since they would indicate the *rate* at which these fragmentation processes occur. Examples, in order of ascending lifetime, are ^3H ($\tau \sim 12 \text{ y}$), ^{14}C ($\tau \sim 5600 \text{ y}$) and ^{10}Be ($\tau \sim 2.5 \cdot 10^6 \text{ y}$).

8'3. *Extra-galactic origin.* — The energy density of the cosmic radiation in the vicinity of the earth is approximately equal to that of star light ($\approx 0.6 \text{ eV/cm}^3$). Thus energetically it is very difficult to accept that an appreciable fraction of the cosmic radiation permeates all space. Instead it is usually assumed that the particles are to a large extent confined to the galaxy, either by magnetic reflection, or by nuclear absorption during diffusion. However, such processes can only be applicable if r , the radius of curvature, is less than the dimensions of the individual magnetic fields in the galaxy. So r must certainly be less than 10^{22} cm , (cf. the galactic radius of $\sim 5 \cdot 10^{22} \text{ cm}$), which in a field of $3 \cdot 10^{-5} \text{ gauss}$, implies that R must be less than 10^9 GV for particles of galactic origin. If the dimensions of homogeneous fields are appreciably less than galactic dimensions, then this limiting rigidity must also be appreciably smaller. The observations of extensive air showers are claimed to have proved the existence of particles with $R \geq 10^9 \text{ GV}$, irrespective of whether

they are protons or heavy nuclei, and show that even these particles do not have any measurable isotropy. Thus it is hardly conceivable that these particles are of galactic origin, and any consideration of the practical difficulty of containing particles in magnetic fields, strongly indicates that an appreciable number of the higher rigidity particles may have originated outside our own galaxy. Consideration of possible accelerating mechanisms also suggest that these nuclei can only attain their vast energies in regions of extra-galactic dimensions.

9. - Future experiments.

At rigidities above about 10^5 GV the extreme rarity of particles imposes the necessity of using the earth's atmosphere as a magnifier of the effects produced by the arrival of each one. Consequently observations have to be made below the atmosphere on the secondary products. At lower rigidities particles can be detected directly, and experiments on the relative abundances and rigidity distributions of the various components must involve detectors carried on space vehicles and very high flying balloons. It is therefore worthwhile considering the characteristics of these carriers in terms of their applicability to specific experiments.

1) Balloons can, or will shortly, carry loads of up to one ton to altitudes where less than 2 g/cm^2 of atmospheric matter lie above the detectors, $\sim 43 \text{ km}$. For many purposes, though not all, this is essentially equivalent to placing the load above the atmosphere. This equipment can generally be recovered for analysis, after a flight whose duration may be as much as $(24 \div 36)$ hours, but which is more usually about 10 hours. During this flight the balloon usually remains closely above its original isocosm, due to the general East-West nature of the atmospheric circulation, so that the threshold rigidity imposed by the geomagnetic field is constant. This feature can be of considerable importance in some experiments. The ascent of the balloon to these very great altitudes takes several hours, which may introduce somewhat uncertain corrections in detectors which have no time resolution, such as emulsions. This slow rise time also limits the usefulness of ballistics in observing transient phenomena, such as flare increases, whose effects may be first detected at sea level, by, for example, neutron monitors or meters.

2) Satellites can, or will shortly, carry loads of several tons to altitudes where they are above the atmosphere and keep them there for a time whose usefulness is generally limited only by the availability of power supplies. The recovery of this load is feasible but not easy, so that telemetry of the data is desirable. The limiting rigidity of those particles able to reach the satellite while it is in orbit almost certainly varies considerably, unless the altitude is

so great that the threshold is very small. Both in principle and practice experiments on satellites are far more complicated and expensive than those on balloons, so that any experiments which can possibly use balloon-carried equipment should do so, in spite of the greater glamour of satellite experiments.

3) Space probes have, for cosmic ray experiments, capabilities similar to satellites except that they can place equipment sufficiently far from the earth that its influences, as a source of magnetic fields and albedo particles, can be neglected. In addition, they can be used to investigate the conditions of appreciable regions of the solar system, which is certain to be of importance in elucidating the nature of the solar modulation process.

Specific experiments may be listed under several headings:

1). *Experiments relating to the charge distribution between the lowest observable rigidities and about 16 GV, the equatorial geomagnetic threshold.* — In this rigidity range the information on the abundances of the various components could still be considerably refined by the use of well established techniques on detectors flown on balloons over various isocosms. Of particular interest are:

a) The charge distribution at very low rigidities, $R < 2.0$ GV. As was seen in Section 4 there is some uncertainty regarding the relative abundance of the various nuclei at these rigidities. These nuclei should be studied in emulsions and counter arrays flown to great altitudes under conditions of low cut-off rigidities, *i.e.*, near the magnetic poles. In particular the relative numbers of protons need to be investigated in detail, preferably using some detector other than the $\check{O}$ -S array so as to check its findings. In addition, the observations that suggest the presence of solar protons at $R < 0.7$ GV, need to be confirmed and extended to other nuclei, although whether such experiments can really be regarded as part of the study of the cosmic radiation seems doubtful.

b) The charge distribution at equatorial cut-off rigidities. The relative abundances are not particularly well known at these rigidities, but are of considerable interest as representing a sample at the highest rigidities where the geomagnetic field provides a threshold. A comparison of relative abundances at these rigidities with those at very low rigidities would provide a sensitive test of any abundance changes, since few of the particles with $R > 2$ GV, say, have $R > 16$ GV. The use of orientated detectors might enable this upper rigidity limit to be raised by a factor of about two as the threshold in the east is greater than that in the vertical direction.

Improved techniques of charge and mass identification are urgently needed in the determination of the abundances of individual elements and isotopes, since current techniques are only just able to resolve mass differences as great as that between ^3He and ^4He , and charge differences below $Z > 10$. The de-

velopment of a nuclear emulsion magnetic spectrometer might provide the necessary mass resolution, although this would require the use of large and powerful magnets and would only be applicable to low rigidity nuclei. The inherent defects of emulsions probably limit the degree of charge resolution obtainable in these detectors, and it will be necessary to develop counter techniques to achieve notably better results. Of special interest in such experiments would be the determination of the abundances of those isotopes or elements, other than L -nuclei, with $Q_z = 0$, such as ^2H , ^3He , ^{10}Be and F . Since some at least of these nuclei will be comparatively rare, and apparatus capable of high precision mass or charge measurements may have small collecting areas, balloons may not provide exposures of sufficient duration; thus imposing the necessity of using satellites. For the emulsion spectrometer these satellites would have to be placed in orbits where they spend a maximum amount of time in regions of low threshold rigidity, and would have to be recoverable. In either case care would have to be taken to see that these orbits did not intersect the radiation belts, which would blacken the emulsions.

Another class of nuclei which should be investigated are those heavier than $\mathcal{VH}$ -nuclei. These, because of their large interaction cross-sections, should be particularly sensitive to the amount of interstellar and interplanetary matter traversed, and measurements of their abundances should help to decide between conflicting theories of the acceleration mechanism. Already their apparent rarity suggests that a process in which heavy nuclei are preferentially accelerated is somewhat unlikely. Because of the low intensity, 10^{-4} that of the $\mathcal{VH}$ -nuclei, if similar to the cosmic abundance, such investigations will have to be done on satellites, which should preferably be in orbits where no nonrelativistic particles can enter the detectors, thus greatly simplifying the determination of charge. If Čerenkov counters are used this would not be necessary, since these impose their own velocity cut-off.

Experiments of the sort described above will have to be made over an appreciable period of time in order to see whether there are any changes in the relative abundances during the solar cycle, as well as to investigate the changes in the rigidity spectra produced by solar modulation. Measurement of these modulation effects appear to offer the best means of studying the processes involved, and thus determining the form of the galactic rigidity spectrum. Observations with space probes of the magnetic field configurations and strengths should also provide information on this subject, suggesting for example, values for the parameters used in Elliott's model. It is conceivable, although perhaps not very likely, that space probes could be used to investigate the galactic radiation directly, by placing detectors outside the solar sphere of influence.

Abundance changes would indicate the presence of solar particles, and might be expected if there is an appreciable solar component having $R > 0.7$ GV,

since it is unlikely that solar-produced particles should traverse the same amounts of matter as those of galactic origin, or even that the source abundances are similar. As an immediate programme, the considerable data obtained during the recent IGY should be compared with similar measurements made during the forthcoming IQSY.

In addition to such long term variations there may also be short term fluctuations of a diurnal or irregular nature. An example of irregular variations are those which occur during flare increases, about which detailed information is still seriously lacking. Similarly some workers, *e.g.* YNGVE [65], though not all, have claimed to have detected appreciable day-night fluctuations in the intensities of slow H -nuclei, which could not, from the sea level data, have been accompanied by similar variations in the intensities of protons or α -particles of the same rigidity. Such variations could occur without being detected at sea level, and might be expected to be of the greatest magnitude for these heavy nuclei, if they are of solar origin, since they are least able to traverse matter without being broken up, and therefore have the smallest probability of being rendered isotropic by any available mechanism. Similarly MEYER [66, 67] has reported variations in $I_{p\alpha}$ during the duration of several balloon flights.

With the advent of satellites and very high flying balloons it is clear that measurements of the charge distributions should not be undertaken in material exposed under more than ~ 5 g/cm² of matter since the uncertainties introduced by the corrections for this amount are excessive by present standards. For example under 5 g/cm² some 30 % of the incident $\mathcal{VH}$ -nuclei have already interacted and must be corrected for. This is particularly true for investigations into the abundances of individual elements or isotopes, as the knowledge of the propagation parameters for such nuclei is extremely sparse, and likely to remain so for some time.

2). *The fragmentation parameters.* — The interpretation of the charge distribution observed near the earth depends on the propagation parameters, λ_i , P_{ij} , in interstellar space and, in the three-dimensional case, on D , the diffusion coefficient. Mean free paths in hydrogen, the predominant interstellar material, can be measured with artificially accelerated particles, while D depends on astronomical evidence. However, the present values of the fragmentation parameters depend on measurements made in emulsions and on radio-chemical analyses of targets bombarded with artificially accelerated protons. Both these methods are open to criticism, and by their nature cannot be capable of giving precise values of these parameters. Preferably measurements will be made with separated emulsions exposed to the primary radiation while immersed in liquid hydrogen. At present several groups, Bristol-St. Louis Bombay, are making measurements in hydrogenous plastics which should pro-

duce appreciably improved values of these parameters in the near future, but more experiments are needed. In particular the question of how these parameters vary with energy needs to be considered in greater detail, over a much greater range of energies than used hitherto.

3). *The charge distribution above 16 GV.* – Above the equatorial threshold rigidity it is necessary to determine not only the charge but the rigidity or energy of each individual particle. In general measurements of the energy are not particularly precise and it is therefore not possible to investigate any detailed structure of the spectra, but only the gross features. In emulsions the characteristics of nuclear interactions may be used to indicate the energy, but, as pointed out previously, due to the small intensities of these energetic particles the detection of such interactions generally depends on at least several hundred GeV being transferred to the electromagnetic cascade, and it is extremely difficult to detect nuclei of intermediate energies. Above $\approx 10^3$ GeV the intensities of X -nuclei are so small that exposures of much greater duration will be needed before reasonable statistics can be obtained. It seems probable that in this region of energies the use of gas filled Čerenkov counters exposed on satellites will make an important contribution. Some idea of the statistical problems of observing these nuclei can be obtained from the fact that an array with a geometric factor of $100 \text{ cm}^2 \cdot \text{sr}$, which is ≈ 10 times, that of the Č-S array and roughly equal to that of a typical emulsion stack, would expect to observe *one* X -nucleus with $R > 10^4$ GV every ten days.

At still higher rigidities observations of the charge distribution can only depend on the development of criteria for distinguishing between extensive air showers induced by individual nucleons and nuclei. For example, MURTHY *et al.* [68] have pointed out that air showers initiated by heavy nuclei should be accompanied by bundles of nucleons in the core of the shower, and that it should be possible to detect these, although, as GREISEN [54] remarks, whether these will be left with high energy after passing through 6 to 12 interaction lengths is doubtful. Alternatively the presence of heavy nuclei might be revealed by measurements of the attenuation lengths of the producing particles made relatively near the top of the atmosphere, *i.e.* under one or two attenuation path lengths from aircraft. A discussion of these problems has been given by GREISEN [54]. Continuing studies of the physical processes occurring during these very high energy interactions should help in the elucidation of the nature of the primary particles.

4). *The form of the energy or rigidity spectra.* – As was shown in Section 6 there are considerable ambiguities in the experimental determinations of the energy spectrum at high energies, even apart from the uncertainty in the composition. Over the whole range there should be a careful search for any

signs of a « break » or sharp change in the form of the spectrum, such as could arise from the superposition of radiation from more than one source, even though experimentally any such discontinuity will only reveal itself as a relatively gradual change in slope. Of particular interest would be the demonstration of a high energy cut-off, indicating as it would a limitation of the acceleration mechanism. The difficulty of such a demonstration is shown by the fact that even if Greisen's spectrum, rather than that of the MIT group, is correct, the number of extensive air showers produced by particles with $U > 10^{10}$ GeV expected to strike an array 1 square km in area is only one every second year. If the MIT spectrum is correct, then such an array would observe one event of this energy every thirteen years.

5). *The spatial distribution.* - Observations on the possible existence of a sidereal variation need to be refined, as the demonstration of a finite effect would provide strong indicative evidence of the nature of the propagation of the cosmic radiation. Such an effect would be expected, *a priori*, to be most prominent for the very high energy, $> 10^9$ GeV, nuclei and observations of high statistical weight are needed on these nuclei. It is possible, however, that such a variation may exist at some intermediate energy, characteristic of the commencement of serious leakage from the galaxy, and that the higher energy particles of extra-galactic origin will show no effect. Thus measurements of variations should be made over the whole range of energies.

* * *

The author is most grateful to all those workers who sent him their results before publication, and in particular to Dr. M. M. SHAPIRO and R. VOGT for sending detailed results prior to their dissemination in preprint form. The assistance of Mrs. P. E. BURTON and Miss E. H. FRENCH, who typed and re-typed this article, and of Mr. W. HARBOR, who prepared the photomicrographs, is gratefully acknowledged.

APPENDIX I

The calculation of intensity ratios.

As was shown in Section 4, if the integral energy spectrum is assumed to be given by eq. (5), then it is possible to derive an expression for the integral rigidity spectrum, eq. (8). For α -particles, this expression for the integral energy spectrum appears to be valid from 0.5 up to at least 10^3 GeV per

nucleon. For the galactic radiation this lower limit may well be less, if $m(E)$, eq. (7), is of only local significance.

Using these equations, expressions can be derived for the ratios of the abundance of various nuclei above a given rigidity, $\Gamma_{ij}(R)$, energy per nucleon, $\Gamma_{ij}(E)$, or total kinetic energy, $\Gamma_{ij}(U)$. These derivations assume that all the nuclei have similar energy spectra, *i.e.* γ is independent of Z .

Then

$$(I.1) \quad \Gamma_{ij}(E) = \frac{\Delta_i}{\Delta_j}$$

and is constant

$$(I.2) \quad \Gamma_{ij}(U) = \Gamma_{ij}(E) \left[\frac{m_0 c^2 + E_j}{m_0 c^2 + (A_j/A_i) E_j} \right]^\gamma$$

so that when $E_j \ll m_0 c^2$, which is the only case when $\Gamma_{ij}(U)$ is of interest

$$(I.3) \quad \Gamma_{ij}(U) = \Gamma_{ij}(E) \left(\frac{A_i}{A_j} \right)^\gamma$$

also

$$(I.4) \quad \Gamma_{ij}(R) = \Gamma_{ij}(E) \left[\frac{(m_0 c^2)^2 + (R)^2 (Z_j/A_j)^2}{(m_0 c^2)^2 + (R)^2 (Z_i/A_i)^2} \right]^{\gamma/2}$$

so that when $R \gg m_0 c^2$

$$(I.5) \quad \Gamma_{ij}(R) = \Gamma_{ij}(E) \left[\frac{Z_j}{Z_i} \cdot \frac{A_i}{A_j} \right]^\gamma$$

and when $R \ll m_0 c^2$

$$(I.6) \quad \Gamma_{ij}(R) = \Gamma_{ij}(E).$$

From eq. (I.4) it can be seen that under these assumptions it is not permissible for $\Gamma_{p\alpha}(R)$ to be constant when R is not much greater than $m_0 c^2$, though it is for all $\Gamma_{ij}(R)$ values where

$$\left[\frac{Z_j A_i}{Z_i A_j} \right] = 1,$$

i.e. essentially for all components other than protons, when it is assumed $A=2Z$.

Experimentally, however, Mc DONALD and WEBBER [2] have presented evidence that $\Gamma_{p\alpha}(R)$ is constant, even for $R < 4.5$ GV, when eq. (I.5) certainly ceases to be a valid approximation. If this observation is to be reconciled with the concept of an integral energy power spectrum then it must be assumed that γ_p , the power index for protons, is a function of energy, since γ_α is known experimentally to be constant.

Then

$$(I.7) \quad \Gamma_{p\alpha}(R) = \frac{\Delta_p}{\Delta_\alpha} \frac{[(m_0 c^2)^2 + (R)^2/4]^{\gamma_\alpha/2}}{[(m_0 c^2)^2 + (R)^2]^{\gamma_p/2}}$$

where Δ_p is a constant, defined in eq. (5) when $\gamma_p = \gamma_\alpha$ and

$$(I.8) \quad \Gamma_{p\alpha}(E) = \frac{\Delta_p}{\Delta_\alpha} (m_0 c^2 + E)^{\gamma_\alpha - \gamma_p}$$

Thus, if $\Gamma_{p\alpha}(R)$ is constant

$$(I.9) \quad \gamma_p = \gamma_\alpha \left[\frac{\log 4[(m_0 c^2)^2 + (R)^2/4]}{\log [(m_0 c^2)^2 + (R)^2]} \right]$$

so that when $R \gg m_0 c^2$

$$(I.10) \quad \gamma_p = \gamma_\alpha$$

and, as before, (I.1),

$$\Gamma_{p\alpha}(E) = \frac{\Delta_p}{\Delta_\alpha}.$$

Obviously, however, eq. (I.9) cannot hold at very low rigidities, as $\gamma_p \rightarrow \infty$ as $(R)^2 \rightarrow 1 - (m_0 c^2)^2$. As an example, when $R = 1$ GV, $\gamma_p = 2.39 \gamma_\alpha$. It is, in any case, inappropriate to consider much lower values of R than 1 GV since the experimental evidence for the constancy of $\Gamma_{p\alpha}(R)$ only extends down to $R > 1$ GV.

In practice $\Gamma_{p\alpha}(E)$ is calculated from an observed proton flux and a calculated α -particle intensity, since the α -particle energy spectrum is better known than that of the protons. Thus, for example, at $R > 4.5$ GV the proton intensity is 610 proton/m² sr s with $E > 3.6$ GeV, while the α -particle intensity is 88 α -particles/m² sr s with $E > 1.5$ GeV per nucleon. From eq. (5), and assuming $\gamma = 1.5$ these figures imply an α -particle intensity of 34.4 α -particles/m² sr s with $E > 3.6$ GeV per nucleon, and thus $\Gamma_{p\alpha}(E) 17.7$ at $E > 3.6$ GeV per nucleon. At $R > 4.5$ GV, $\gamma_p = 1.04 \gamma_\alpha$, (I.9), and so $\Gamma_{p\alpha}(E)$ at high energies is, from (I.8) equal to 18.4.

$\Gamma_{p\alpha}(E)$ at high energies can also be determined from (I.5) if $\Gamma_{p\alpha}(R)$ is known. Assuming $\Gamma_{p\alpha} = 7.0 \pm 0.4$ for results where R is not too small, $\Gamma_{p\alpha}(E) = 19.8 \pm 1.2$ if all the α -particles are ⁴He-nuclei, but will be less if there are ³He-nuclei as well.

A completely alternative approach is provided by the experimental fact that the data on the intensities of α -particles can equally well be expressed as an integral rigidity spectrum of the form

$$(I.11) \quad J_\alpha(R) = \frac{A}{R^\nu}.$$

Figure 16 shows how good a fit this equation is for $2.0 < R < 15$ GV. and that it is no way inferior to the derived integral rigidity spectrum, eq. (8). From this figure, $\nu \approx 1.15$, which suggests that this spectrum will not fit the data at higher rigidities. However, the use of eq. (I.11) leads to some interesting results.

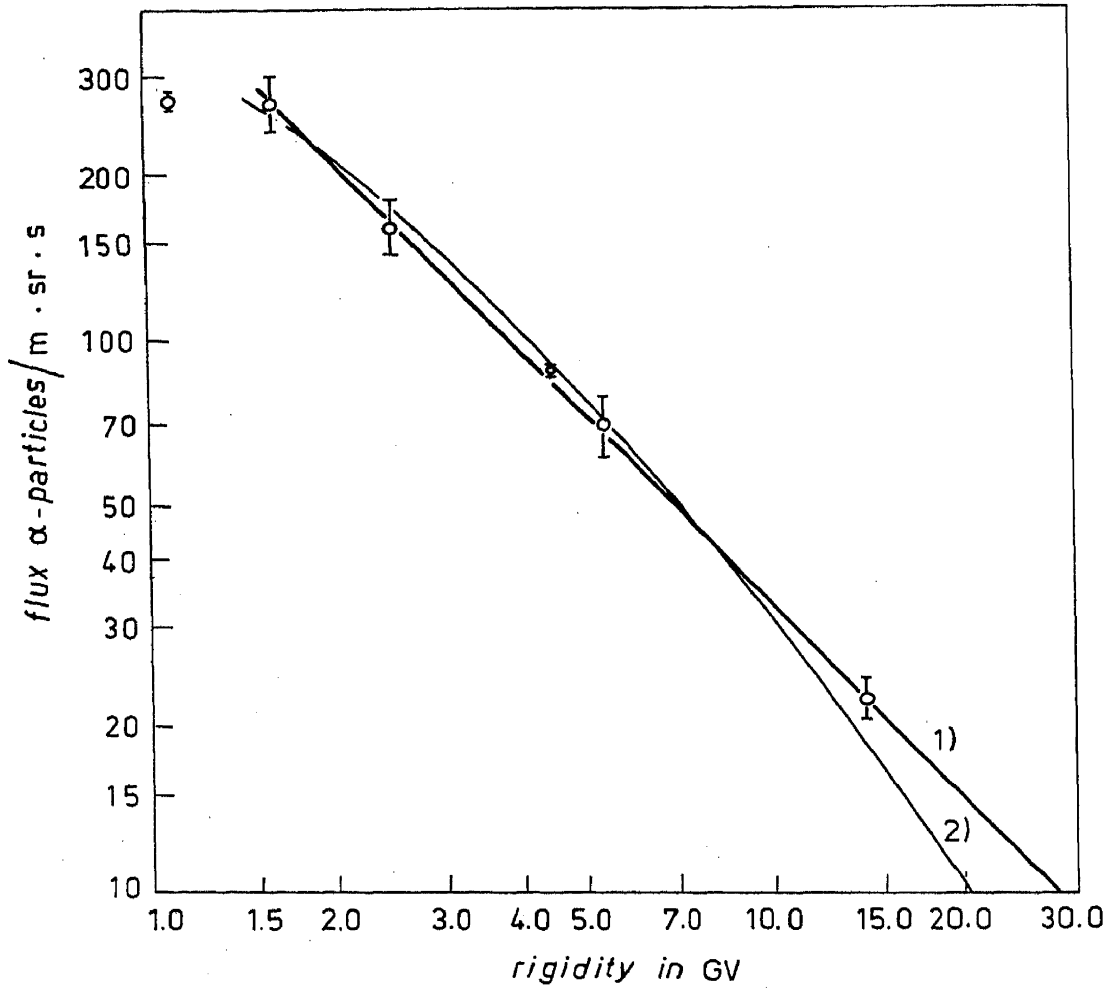


Fig. 16. - The integral rigidity spectrum of α -particles using intensities taken from WADDINGTON [8]. Curve 1 shows eq. (I.11), while curve 2 shows eq. (8), the derived integral rigidity spectrum.

From eq. (I.11) it is possible to derive an integral energy spectrum of the form

$$(I.12) \quad J_{\alpha}(E) = \frac{(Z_{\alpha}/A_{\alpha})A}{[E(E + 2m_0c^2)]^{1/2}}$$

and from (I.11) and (I.12) it appears that

$$(I.13) \quad \Gamma_{ij}(R) = \frac{A_i}{A_j} \quad \text{and is constant,}$$

$$(I.14) \quad \Gamma_{ij}(E) = \Gamma_{ij}(R) \left[\frac{Z_i}{Z_j} \cdot \frac{A_j}{A_i} \right]^{\nu} \quad \text{and is constant if } \nu \text{ is}$$

and

$$(I.15) \quad \Gamma_{ij}(U) = \Gamma_{ij}(E) \left[\frac{E_j(E_j + 2m_0c^2)}{(A_j/A_i)E_j[(A_j/A_i)E_j + 2m_0c^2]} \right]^{\nu/2}$$

so that when $E \gg m_0c^2$

$$(I.16) \quad \Gamma_{ij}(U) = \Gamma_{ij}(E) \left(\frac{A_i}{A_j} \right)^{\nu} \quad \text{and is constant.}$$

Thus on this approach, both $\Gamma_{ij}(R)$ and $\Gamma_{ij}(E)$ appear to be constants, and there is thus no necessity to assume ν varies sharply at low R . However, from a comparison of (I.11) with eq. (5), the integral energy spectrum, it can be seen that when $E \gg m_0 c^2$

$$(I.17) \quad \gamma = \nu$$

and so ν must vary between $R \approx 15$ GV where it is ≈ 1.15 , and higher rigidities, where it is 1.5. Consequently $\Gamma_{ij}(E)$ is variable after all, as it must be from the energy spectrum analysis.

The integral rigidity spectrum is thus not as good a fit to the data over as wide a range of rigidities as is that spectrum derived from the integral energy spectrum. Therefore, this integral energy spectrum still appears to be more basic in its description of the physical situation, although its form for the galactic radiation is not well determined at low energies.

APPENDIX II

The effect of ^3He -nuclei on the α -particle energy spectra.

Experimentally the quantities observed are:

i) Residual range in emulsion, ϱ . Here the kinetic energy per nucleon, E is given to a reasonable approximation by

$$(II.1) \quad E = 0.25 \left(\frac{Z^2}{A} \cdot \varrho \right)^{0.58},$$

where E is in MeV per nucleon and ϱ in microns.

As a consequence a ^3He nucleus which is erroneously assumed to be a ^4He -nucleus, will have its energy *underestimated* by a factor of $(4/3)^{0.58}$.

ii) The multiple Coulomb scattering in emulsion. Here the quantity $p\beta$, where p is the momentum and β is the particle velocity, is given by

$$(II.2) \quad p\beta = ZK/\bar{\alpha},$$

where K is the scattering constant and $\bar{\alpha}$ the scattering parameter. Thus if p_N is the momentum per nucleon

$$(II.3) \quad p_N\beta = \frac{Z}{A} \cdot \frac{K}{\bar{\alpha}}.$$

As a consequence a ^3He nucleus will have a $p_N\beta$ value which is *underestimated* by a factor of $4/3$. Over the range of energies, $(100 \div 700)$ MeV per

nucleon, of interest, E is approximately proportional to $p_N\beta$, and so the error in E is also about $4/3$.

iii) The rate of energy loss, dE/dx , in a scintillator or emulsion. Here

$$(II.4) \quad dE/dx \propto Z^2/\beta^2$$

in the range of energies of interest. Thus a ${}^3\text{He}$ nucleus will have the same dE/dx as a ${}^4\text{He}$ nucleus when it has the same velocity as the ${}^4\text{He}$ nucleus, *i.e.* when it has $4/3$ times the energy. So the energies of ${}^3\text{He}$ nuclei are *under-*estimated by a factor of $4/3$ but the energies per nucleon are correct.

Thus all the common experimental methods of energy determination lead to underestimates of the true values for ${}^3\text{He}$ nuclei. The correction that should be applied to the experimentally determined energy spectra can be calculated, if it is assumed that the ratio

$$(II.5) \quad \psi = \frac{{}^4\text{He}}{{}^3\text{He} + {}^4\text{He}}$$

is a constant independent of energy over the narrow range of energies involved. Then if $dJ/dE = f(E)$ is the function observed assuming that all the α -particles are ${}^4\text{He}$ nuclei, and σ is the error made in the energy estimate by this assumption,

$$(II.6) \quad \left(\frac{dJ_4}{dE}\right)_{E_1} = \psi \cdot f(E_1)$$

and

$$(II.7) \quad \left(\frac{dJ_3}{dE}\right)_{E_1} = (1 - \psi) \cdot f(E_1),$$

where J_4 is the intensity of the ${}^4\text{He}$ nuclei and J_3 is the intensity of ${}^3\text{He}$ nuclei. Hence

$$(II.8) \quad \left(\frac{dJ_{4+3}}{dE}\right)_{\sigma E_1} = (1 - \psi) \cdot f(E_1) + \psi \cdot f(\sigma E_1)$$

is the true differential energy spectrum.

Figure 10 shows $f(E)$, eq. (6), together with the resultant value of dJ_{4+3}/dE when $\psi = \frac{1}{2}$. The correction can be seen to be rather small, in any case appreciably less than the statistical uncertainties of the experimental points. It should be noticed that these points have not been corrected for any ${}^3\text{He}$ present, and have been determined by various experimental techniques.

These uncertainties can be almost completely removed by expressing the results as a rigidity spectrum, since in calculating rigidities values from experimental energy or momentum values the rigidities of the ${}^3\text{He}$ nuclei are *over-*estimated by a factor of $4/3$, this largely cancelling the previous underestimate of the energy or momentum.

REFERENCES

- [1] F. B. McDONALD: *Phys. Rev.*, **104**, 1732 (1956).
- [2] F. B. McDONALD and W. R. WEBBER: *Phys. Rev.*, **115**, 194 (1959).
- [3] C. J. WADDINGTON: *Phil. Mag.*, **5**, 1105 (1960).
- [4] K. KRISTIANSSON, O. MATHIESEN and B. WALDESKOG: *Arkiv. f. Fysik*, **17**, 427 (1960).
- [5] M. WADA: *Journ. Sci. Res. Inst.*, **51**, 201 (1957).
- [6] R. R. DANIEL, P. J. LAVAKARE and P. K. ADITYA: *Nuovo Cimento*, **17**, 837 (1960).
- [7] S. HAYAKAWA, K. ITO and Y. TERASHIMA: *Prog. Theor. Phys. Suppl.*, **6**, 1 (1958).
- [8] C. J. WADDINGTON: *Prog. Nucl. Phys.*, **8**, 1 (1960).
- [9] W. R. WEBBER: *Prog. in Cosmic Ray Phys.*, **6**, 77 (1961).
- [10] P. S. FREIER, E. P. NEY and C. J. WADDINGTON: *Phys. Rev.*, **113**, 921 (1959).
- [11] F. W. O'DELL, M. M. SHAPIRO and B. STILLER: private communication see also *Proc. Moscow Conf.*, **3**, 118 (1961), and *Journ. Phys. Soc. Japan*, **17**, Suppl. AIII, 23 (1962).
- [12] I. J. VAN HEERDEN and B. JUDEK: *Can. Journ. Phys.*, **38**, 964 (1960).
- [13] C. J. WADDINGTON: *Phil. Mag.*, **2**, 1059 (1957).
- [14] M. KOSHIBA, G. SCHULTZ and M. SCHEIN: *Nuovo Cimento*, **9**, 1 (1958).
- [15] C. M. GARELLI, B. QUASSIATI and M. VIGONE: *Nuovo Cimento*, **15**, 121 (1960).
- [16] M. V. K. APPA RAO, S. BISWAS, R. R. DANIEL, K. A. NEELAKANTAN and B. PETERS: *Phys. Rev.*, **110**, 751 (1958).
- [17] R. R. DANIEL and N. DURGAPRASAD: *Suppl. Nuovo Cimento*, **23**, 82 (1962).
- [18] C. Q. ORSINI: *Nuovo Cimento*, **16**, 1040 (1960).
- [19] R. CESTER, A. DEBENEDETTI, C. M. GARELLI, B. QUASSIATI, L. TALLONE and M. VIGONE: *Nuovo Cimento*, **7**, 371 (1958).
- [20] A. ENGLER, M. F. KAPLON and J. KLARMANN: *Phys. Rev.*, **112**, 597 (1958).
- [21] H. E. SUESS and H. C. UREY: *Rev. Mod. Phys.*, **28**, 53 (1956).
- [22] A. G. W. CAMERON: *Ap. Journ.*, **129**, 672 (1959).
- [23] F. B. McDONALD: *Phys. Rev.*, **116**, 462 (1959).
- [24] P. H. FOWLER, P. S. FREIER and E. P. NEY: *Suppl. Nuovo Cimento*, **8**, 492 (1958).
- [25] F. B. McDONALD: *Phys. Rev.*, **104**, 1723 (1956).
- [26] A. ENGLER, F. FOSTER, T. L. GREEN and J. MULVEY: *Nuovo Cimento*, **20**, 1157 (1961).
- [27] H. ELLIOT: *Phil. Mag.*, **5**, 54 (1960), and these proceedings.
- [28] H. AIZU, Y. FUJIMOTO, S. HASEGAWA, M. KOSHIBA, I. MITO, J. NISHIMURA K. YOKIO and M. SCHEIN: *Phys. Rev.*, **121**, 1206 (1961).
- [29] F. B. McDONALD and W. R. WEBBER: *Journ. Geo. Res.*, **65**, 767 (1960).
- [30] M. V. K. APPA RAO: *Phys. Rev.*, **123**, 295 (1961).
- [31] H. V. NEHER: *Phys. Rev.*, **107**, 588 (1957).
- [32] J. R. WINCKLER and K. A. ANDERSON: *Phys. Rev.*, **108**, 148 (1957).
- [33] R. VOGT: *Bull. Amer. Phys. Soc.*, **6**, 263 (1961); *Phys. Rev.*, **125**, 366 (1962).
- [34] D. EVANS: *Proc. Moscow Conf.*, **3**, 92 (1959).
- [35] E. TAMAI: *Phys. Rev.*, **117**, 1345 (1960).
- [36] Y. TERASHIMA: *Prog. Theor. Phys.*, **23**, 1138 (1960).
- [37] C. E. FICHTEL: *Nuovo Cimento*, **19**, 1100 (1961).
- [38] C. L. CRITCHFIELD, E. P. NEY and S. OLEKSA: *Phys. Rev.*, **85**, 461 (1952).

- [39] P. MEYER and R. VOGT: *Phys. Rev. Lett.*, **6**, 193 (1961).
- [40] J. A. EARL: *Phys. Rev. Lett.*, **6**, 125 (1961).
- [41] J. DUTHIE, C. M. FISHER, P. H. FOWLER, A. KADDOURA, D. H. PERKINS, K. PINKAU and W. WALTER: *Phil. Mag.*, **6**, 89 (1961).
- [42] V. L. GINZBURG: *Prog. in Cosmic Ray Phys.*, **4**, 337 (1958).
- [43] A. ENGLER and U. HABER-SCHAIM: *Phys. Rev.*, **95**, 1700 (1954).
- [44] D. LAL: *Proc. Ind. Acad. Sci.*, A **38**, 93 (1953).
- [45] M. F. KAPLON, B. PETERS, H. L. REYNOLDS and D. M. RITSON: *Phys. Rev.*, **85**, 295 (1952).
- [46] P. H. FOWLER and C. J. WADDINGTON: *Phil. Mag.*, **1**, 637 (1956).
- [47] E. LOHRMANN and M. W. TEUCHER: *Phys. Rev.*, **115**, 638 (1959).
- [48] P. L. JAIN, E. LOHRMANN and M. W. TEUCHER: *Phys. Rev.*, **115**, 654 (1959).
- [49] H. ALY, J. DUTHIE, A. KADDOURA, D. H. PERKINS and P. H. FOWLER: *Proc. 1960, Rochester Conf.*, p. 829, and private communication.
- [50] P. H. BARRETT, L. M. BOLLINGER, G. COCCONI, Y. EISENBERG and K. GREISEN: *Rev. Mod. Phys.*, **24**, 133 (1952).
- [51] J. DUTHIE, P. H. FOWLER, A. KADDOURA, D. H. PERKINS, K. PINKAU and W. WALTER: see [41].
- [52] K. GREISEN: *Prog. in Cosmic Ray Phys.*, **3**, 1 (1956).
- [53] G. COCCONI: *Suppl. Nuovo Cimento*, **8**, 560 (1958).
- [54] K. GREISEN: *Ann. Rev. Nucl. Sci.*, **10**, 63 (1960).
- [55] G. W. CLARK, J. EARL, W. L. KRAUSHAAR, J. D. LINSLEY, B. B. ROSSI, F. SCHERB and D. W. SCOTT: *Phys. Rev.*, **122**, 637 (1961) and private communication (1961).
- [56] J. C. BARTON: private communication (1961); see also J. DUTHIE, P. H. FOWLER, A. KADDOURA, D. H. PERKINS, K. PINKAU and W. WALTER [41, 51].
- [57] A. E. CHUDAKOV, N. M. NESTEROVA, V. I. ZATSEPIN and E. I. TUKISH: *Proc. Moscow Conf.*, **2**, 50 (1959).
- [58] B. PETERS: *Proc. Moscow Conf.*, **3**, 157 (1959).
- [59] L. I. DORMAN: *Proc. Moscow Conf.*, **4**, 7 (1959).
- [60] L. I. DORMAN: *Cosmic Ray Variation* (American trans.) (Moscow, 1957).
- [61] C. E. FICHEL and D. E. GUSS: *Phys. Rev. Lett.*, **6**, 495 (1961).
- [62] C. Y. FAN, P. MEYER and J. A. SIMPSON: *Phys. Rev. Lett.*, **5**, 272 (1960).
- [63] V. L. GINZBURG and S. I. SYROVAT-SKII: *Sov. Phys. Uspekhi*, **3**, 504 (1961).
- [64] F. HOYLE: *Month. Not. Roy. Astr. Soc.*, **120**, 338 (1960).
- [65] V. H. YNGVE: *Phys. Rev.*, **92**, 428 (1953).
- [66] P. MEYER: *Phys. Rev.*, **115**, 1734 (1959).
- [67] P. MEYER: *Journ. Geo. Res.*, **65**, 3881 (1960).
- [68] G. T. MURTHY, B. PETERS, P. V. RAMANAMURTHY and B. V. SHREEKANTAN: *Proc. Moscow Conf.*, **3**, 150 (1959).

Use of Solid-State Detectors as Energy Spectrometers for Cosmic High-Energy Particles.

J. LABEYRIE

Service d'Electronique Physique, Centre d'Etudes Nucléaires de Saclay - Saclay

1. - General properties.

Three or four years ago a new type of nuclear particle detector was developed: the solid-state junction detector.

At the present time it looks like a flat cylinder, with a surface about one square centimeter, and a height which has an order of magnitude of one millimeter.

The material is a semi-conductor of high resistivity, silicon in most cases.

On each flat side, an electrode is deposited, and a difference of potential of about 100 V is applied between these electrodes. When a nuclear ionizing particle crosses this silicon piece, an electric pulse appears between the two electrodes.

This device acts as an ionization gaz chamber. It may be used as such, *i.e.* as a detector of nuclear particles, it may be also used to measure the energy lost by the particle in the detector, by measuring the amplitude of the pulse produced.

One can find in recent litterature many papers on the properties of this new kind of detectors, and, also, description of how to make them.

Last year a number of papers illustrating the state of development of the technique was collected in the *I.R.E. Proceedings of the Gatlinburg Congress* [1].

Several advantages appear, at once, over the ionization gas chamber or counter:

- 1) Being a solid, the density of matter is about one thousand times as great as in a gas chamber or counter, and it follows that by unit length of path

of the particle the number of ions produced would be a thousand times higher, if the mean ionization per ion pair loss were the same.

2) But, this loss is about ten times lower in solid state detectors than in gas detectors (for instance only 3.55 eV are needed to produce one ion pair in silicon, instead of about 35 eV in air) and consequently, the number of ions produced per unit path length of the particle trajectory will be about ten thousand times greater than in a gas.

This has enabled people to plot beautiful α energy spectra, where lines are about three times narrower than those obtained with gas chambers.

3) There is no « wall effect » as in gas detectors: the solid state detector may be surrounded by air or may be, as well, in vacuum, and by this way, many parasitic effects are absent.

4) The duration of the pulse is very short: 10^{-9} s is a frequent order of magnitude. In very thin solid state detectors, half of the whole amplitude of the pulse can be obtained in less than 10^{-10} s. This may be an advantage when one has to fight against numerous background pulses.

Until recently however, it was not possible to use these detectors for detecting particles producing a very small ionization, such as γ -rays or very fast particles which are in vicinity of the ionization minimum. This was due to the fact that these solid state detectors had only a very small thickness, about fifty microns, and that they had also a parasitic current between the two electrodes: the possibility of detection failed when the current due to the ionization process had about the same value as that of the parasitic current.

2. - High-energy particles detection.

Now, things have changed and we are able to make crystals of big thickness (several millimeters), and of very high resistivity; this last result has been obtained by compensating impurities which gave parasitic current by specially added atoms; lithium, for example, is a good compensator for highly purified commercial silicon.

I will give now some results obtained in the field of energy spectrometry of very energetic particles, such as protons or mesons in the 0.5 to 1.5 GeV range of energy.

First results have been given by MILLER, FOREMAN, YUAN [2], of Brookhaven National Laboratory, by VAN PUTTEN, VANDER VELDE [3] at University of Michigan, and by KOCH, MESSIER, VALIN [4] at Saclay. Very recently the

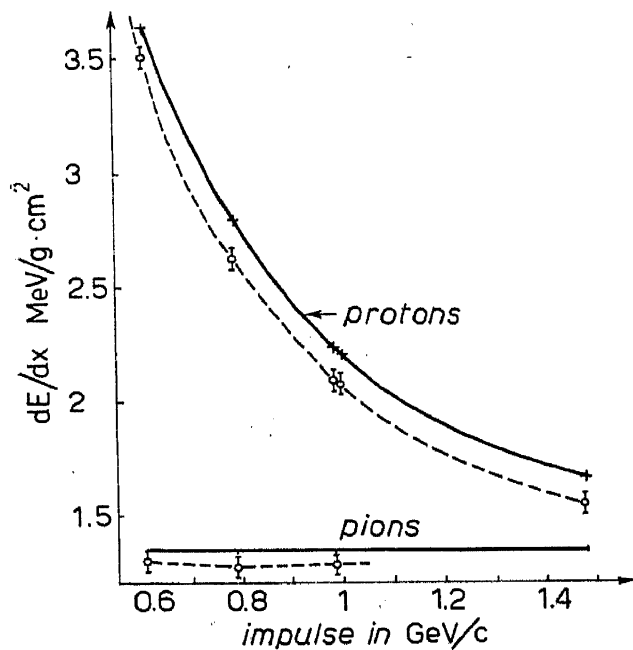


Fig. 1. - dE/dx vs. momentum for pions and protons in silicon: $\odot$ experimental, ——— calculated curve from Rossi's formula, with «density corrections» for silicon.

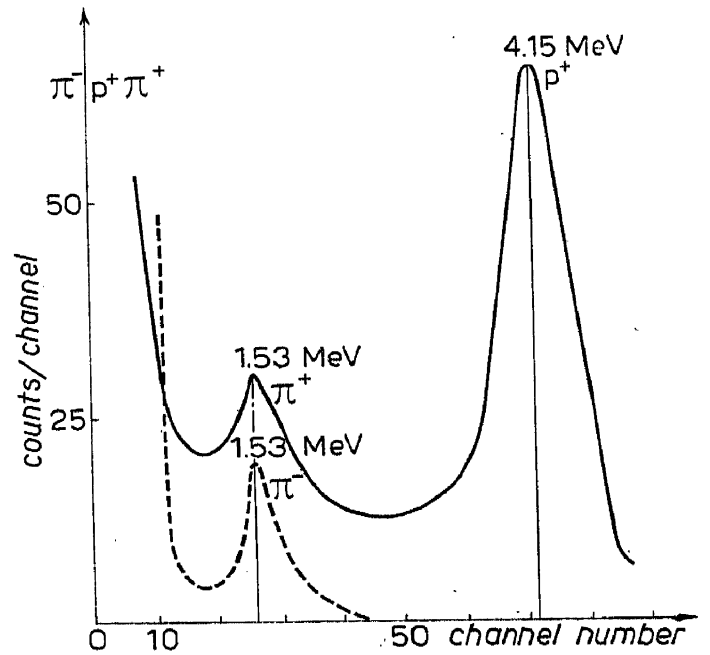


Fig. 2. - Line spectrum of pions and protons of 0.615 GeV/c, NIP silicon diode, $e = 4.93$ mm parallelly to the junction plane.

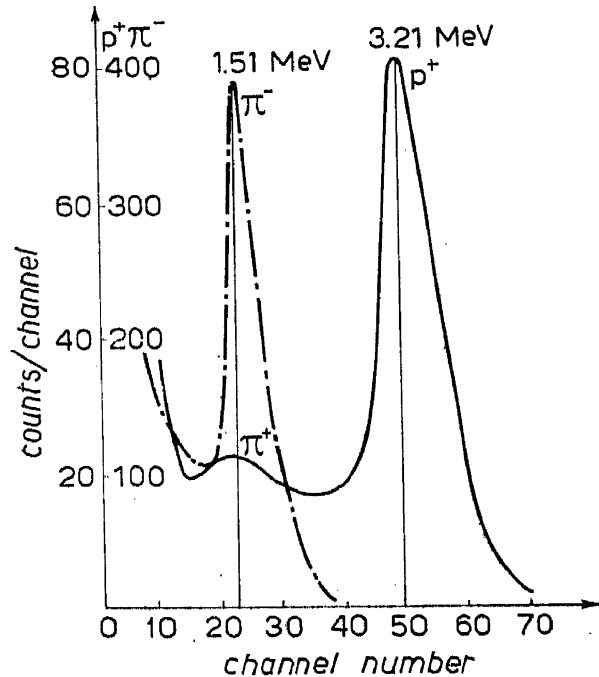


Fig. 3. - Line spectrum of pions and protons of 0.787 GeV/c, NIP silicon diode, $e = 4.93$ mm parallelly to the junction plane,

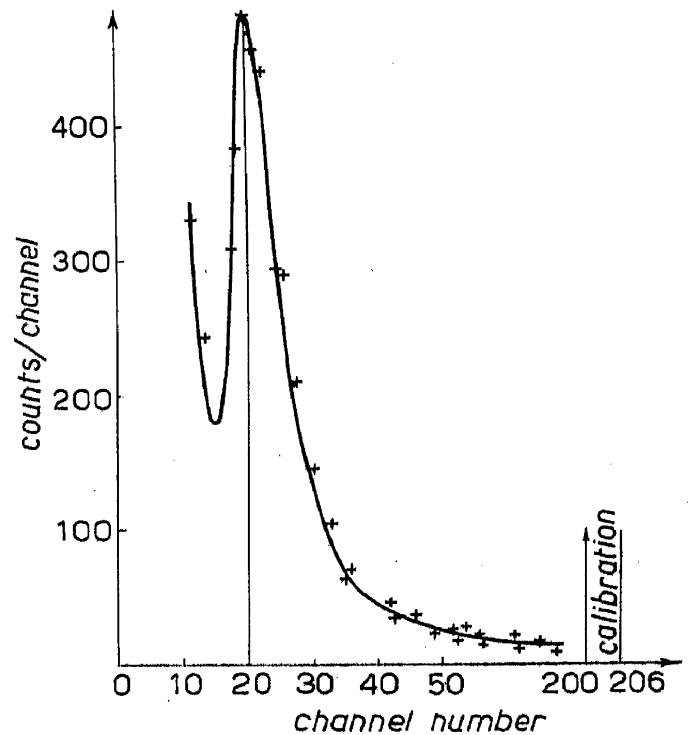


Fig. 4. - Spectrum line of protons of 1.0 GeV/c, lithium-silicon diode, $e = 0.23$ mm.

latter authors have obtained, at Saclay, the specific loss of energy of relativistic mesons and protons as a function of momentum, for silicon.

The Fig. 1 shows this curve: between 0.6 and 1.5 GeV/c, the experimental specific dE/dx expressed in MeV/g/cm² decreases from 4.0 to 1.6 for protons. In the same range of energy, the pions + and -, which are at minimum ionization give a constant dE/dx , equal to 1.27 MeV/g/cm².

On this curve KOCH *et al.* have plotted theoretical values obtained by taking in account the density effect, calculated from STERNHEIMER [5] and with an ionization potential of 179.5 eV. There is a small constant difference of about 4% between theoretical and experimental curve, which has not been completely explained till now.

In Fig. 2 and 3, we can see typical spectral lines obtained for monoenergetic lines of protons and pions having the same momentum, respectively 615 MeV/c and 787 MeV/c.

These curves have been obtained with « NIP » junctions where the electrodes were respectively boron-doped and phosphorus-doped silicon, and the sensitive part was highly pure silicon of resistivity 24 000 ohm-cm. The thickness crossed by the particle was 4.93 mm.

On Fig. 2 one can see, for example, that the energy lost by 615 MeV/c momentum particles in this silicon thickness is 1.53 MeV for pions and 4.15 MeV for protons.

On Fig. 4 one can see the peak due to protons having a momentum of 1.0 GeV/c, crossing only 0.23 mm of such a silicon detector (*): it loses sufficient energy, however, (0.127 MeV) to give a very good measurement of the momentum of these protons, using the plot of Fig. 1.

The peak width at half-height is about 20% at 1 GeV, which corresponds to the Landau effect for a 1 GeV proton losing 0.127 MeV.

3. - Precision of energy determination for protons at minimum ionization.

At an energy corresponding to minimum ionization, particles with one electric charge lose 315 keV/mm of Si, or 1.30 MeV/g/cm² giving 90 000 ion pairs per mm of Si.

In a typical diode of 1 mm thickness having a capacity of about 10 pF/cm², the impulse height obtained by a particle at minimum ionization per mm of path in silicon is $1.44 \cdot 10^{-3}$ V.

As in all ionization devices, the energy discrimination is not possible in the vicinity of energy corresponding to minimum ionization.

(*) Density: 2.4 g cm⁻³.

But for energy smaller than that corresponding to minimum ionization the energy loss per mm of Si increases very quickly when energy is decreasing, and the energy discrimination becomes very precise.

As we can see from the table, for example, one proton having a momentum of 1000 GeV/c loses about 0.5 MeV/mm and produces 140 000 ion pair/mm, and a proton having a momentum of 1.478 GeV/c loses 0.380 MeV/mm and produces 108 000 ion pair/mm.

TABLE I. - *Energy loss and ion pairs produced in 1 mm of Si.*

Range (mm)	$p\beta$ (MeV/c)	π mesons				protons			
		MeV	MeV/mm	MeV/g cm ²	Ni/mm	MeV	MeV/mm	MeV/g cm ²	Ni/mm
4.93	615	1.53	0.315	1.30	90.000	4.15	0.84	3.51	236.000
4.93	787	1.51	0.310	1.27	85.000	3.21	0.65	2.62	183.000
4.93	984	1.52	0.312	1.28	86.000	2.45	0.5	2.09	140.000
0.23	1 000					2.43	0.495	2.08	139.000
4.93	1 478					1.80	0.378	1.54	106.000

If we use an amplifier having a rise time of about 1 μ s and equivalent background noise level of 13 keV, the overall background noise level of these diodes+amplifier is about 16 keV, for a diode thickness of 0.3 mm having 0.5 cm² surface. With this diode we can discriminate at limit two ionization processes giving a difference of about 20 keV, or about 5 700 ion pairs. This would correspond to a precision in energy discrimination which is about 8 MeV or 0.8 % for protons having an energy of about 1 GeV.

However, due to the Landau effect, it is possible to separate two lines in 1 GeV region of the proton energy spectrum only if they are separated by about 5 to 10 % (or 50 to 100 MeV) for a 5 mm diode.

4. - Discrimination between two relativistic particles of different Z .

It we now consider the problem of how to distinguish the Z number of two relativistic particles having energies above the minimum of ionization, it seems that the solid state detector presents some advantage over other devices.

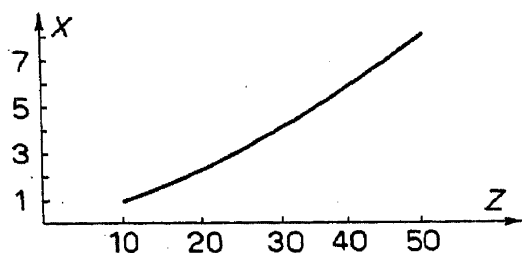


Fig. 5. — On the y axis the number X such as it is possible to discriminate between Z and $Z - X$.

If it would be possible to measure the total energy loss in the detector, calculations made by LAUZANNE and MESSIER [6] show that with one silicon detector of 2 mm thickness giving dE/dx for particles of different Z having energy from 1 to 100 GeV per nucleon it would be possible to discriminate Z , as we can see on Fig. 5.

If only a few per cent of energy loss is measured by the detector (as is the case with nuclear emulsions) the problem of Z selection is more complicated. Calculations are now in progress.

Due to theoretical analysis of Landau's effect several coincident detectors will probably give a better measurement of the most probable value of dE/dx and will probably allow selection of the Z 's of primary particles at ionization minimum.

5. — Saturation.

Measurements of the height of pulses due to fission particles, using « NP » diodes [7] have shown that the saturation effects do not appear till at least an ionization loss of 30 MeV/mg cm².

As a singly charged particle loses 1.3 keV/mg cm², that means that the solid state counter may be crossed by particles losing 23 000 times more energy than singly charged particles without showing saturation effects. That means that we could observe, without saturation, particles having $Z = 150$.

6. — Surface of detection.

It is possible to detect heavy relativistic particles, carbon ions for example, with a surface of about 2 cm², with only one preamplifier (*). Relativistic pions give about 300 keV/mm and a relativistic carbon ion will give about 36 times more, or 12 MeV. The background noise is 60 keV/cm² and will be about 3 MeV for 30 cm² diode. The ratio of signal on noise will be about 200 times smaller as for a one cm² diode. We can see from above that the 12 MeV ionization loss due to carbon ions will be the limit of detection, giving a signal on noise ratio equal to unity with a 30 cm² diode.

(*) Charge preamplifier where the output signal is proportional to the quantity of charge lost in the detector and is independent of sufficiently low input capacitance.

REFERENCES

- [1] *Proceedings of the Seventh Annual Meeting*, in *I.R.E. Trans.*, **8**, 1 (1961).
- [2] G. L. MILLER, B. M. FOREMAN and L. C. L. YUAN: *I.R.E. Trans.*, **8**, 1 (1961).
- [3] J. D. VAN PUTTEN and J. C. VANDER VELDE: *I.R.E. Trans.*, **8**, 1 (1961).
- [4] L. KOCH, J. MESSIER and J. VALIN: *Ind. International Symposium on Nuclear Electronics* (Beograd, May 1961).
- [5] R. M. STERNHEIMER: *Phys. Rev.*, **115**, 137 (1959).
- [6] C. LAUZANNE and J. MESSIER: private communication.
- [7] L. KOCH, J. MESSIER and J. VALIN: *Compt. Rend. Acad. Sci.*, **251**, 2912 (1960).

Extensive Air Showers.

L. SCARSI

Istituto di Scienze Fisiche dell'Università - Milano
Istituto Nazionale di Fisica Nucleare - Sezione di Milano

Introduction.

At the time of their discovery, the Extensive Air Showers were attributed to electromagnetic cascades of primary electrons or photons. Not until several years later were they correctly correlated with the chain of events following the interaction in the earth's atmosphere of the primary nucleonic component. The common destiny of primary protons or heavy nuclei, interacting with the nuclei of the atmosphere, is to produce a shower of particles, whose development is dominated by the nuclear-active component. The depth at which the shower will reach its maximum, and its size at different stages, is uniquely determined by the energy of the primary. The energy and lateral distribution of the shower electromagnetic component, which constitutes the bulk of the particles outside the «core», is mainly due to scattering and relatively low energy processes. If the energy of the initiating primary particle is high enough, the shower will reach the ground as a conspicuous phenomenon; usually we refer as E.A.S. to showers with a total of $\geq 10^5$ charged particles at sea level (corresponding to a primary energy of $\geq 10^5$ eV.)

There are two main interests in investigating Extensive Air Showers. The first, more direct, to gain information on the high energy interactions in a region beyond the limit which it is possible to reach with accelerators ($\sim 10^{10}$ eV at present and $\sim 10^{11}$ eV as a best hope for the future), or with detectors flown at the top of the atmosphere ($(10^{13} \div 10^{14})$ eV). This requires a very detailed knowledge of the shower structure, and consequently a complex system of detection able to separate the various components and to determine their energy distribution, eventually at successive stages of the shower development. Such a system cannot be very large, and since events fall off rapidly with energy, results to date have been obtained only up to $\sim 10^{17}$ eV.

The second interest is related to the investigation of the far energy spectrum of the primary Cosmic Rays, their direction of arrival from the Celestial Sphere, their composition and so to the general problem of the origin of the Cosmic Radiation. In this case, E.A.S. are used mainly as a technique of detection, more than for the sake of the phenomenon in itself. Counting rate requirements, and so sensitive area covered, is of primary importance; detectors are simpler and the relative spacing greater than in the preceding case. For individual events, the size is measured with less accuracy, sometimes within a factor of two. On the other hand high precision is not, at least for the moment, essential in a field in which many parameters are still undetermined to an order of magnitude.

In this Seminar, I will describe some of the aspects of the technique of detection of large E.A.S., referring especially to the M.I.T. installations at Volcano Ranch (New Mexico - U.S.A.). Unfortunately I will be able to anticipate only few of the results obtained with this array, the data being still in course of processing. Preliminary results will be communicated next September at the International Conference on Cosmic Rays in Kyoto.

1. - E.A.S. and the cosmic ray primary energy spectrum.

Experimental information on the primary Cosmic Radiation, derived by the direct detection of the primary particles themselves, does not go beyond about 10^{13} eV. The value of the intensity for the incoming flux ($J(E) \geq 10^{13}$ eV) $\sim 10^{-6}$ s $^{-1}$ cm $^{-2}$ sr $^{-1}$) and the steepness of the integral spectrum ($J \propto E^{-2}$) put out of range detectors of a size which is reasonable to put at the top of the atmosphere or outside it. To reach higher energies we have to turn to some kind of geometrical amplifier to compensate for the small intensity with a large sensitive area. Furthermore, it will help substantially to subdivide the incoming energy in as many as possible secondary particles in order to make a calorimetric type of measurement of the energy released. It is exactly what the atmosphere does through the E.A.S., allowing us to make our measurements on the ground, with no limitations of time, weight and space, if not of money to invest in detectors and people to run them.

2. - The M.I.T. « E.A.S. » program.

In the last ten years, investigation of the E.A.S. has been one of the main projects of the M.I.T. Cosmic Ray group directed by BRUNO ROSSI. The primary purpose of the experiments, developed in successive steps, has been the study of the primary Radiation for the sake of information which it is possible to obtain about the problem of the origin. The secondary purpose, neces-

sary on the other hand to cover the first, has been the analysis of the structure of the showers themselves.

The line followed is the one of obtaining as much complete information as possible about each shower detected. The experimental method is a combination of «fast timing» and «density sampling». The first is based on the fact that the shower front is rather well defined, the bulk of the particles being concentrated in a thin disc normal to the axis and travelling along the axis with nearly the speed of light. The direction of the axis, and of the primary, is determined by the relative arrival time of the shower particles at several widely separated detectors. The «density sampling» is based on the fact that the particle density in a shower is radially symmetric and decreases monotonically with distance from the axis. This allows the reconstruction of the shower by analysing the pattern of particle densities registered by the same detectors used for timing.

The ideal array of detectors would cover the complete area struck by particles belonging to the shower and be subdivided in an almost infinite number of sections. In practice, for sizes above 10^6 particles at sea level, this is never possible, the total number of detectors and their surface being limited by practical and technical reasons. In planning an array of detection, there are two determining factors:

- a) the range of sizes to be investigated,
- b) the expected counting rate in this range.

Having already some information on the lateral distribution of the particles on the shower front by extrapolation from preceding experiments (and this establishes a natural line of experiments, from smaller and more numerous showers to bigger and less frequent), 5 to 10 points with measured density

TABLE

Experiment	Location	Type of detector	Array
Agassiz [1]	Harvard (Mass.) 180 m.a.s.l.	$0.9 \text{ m}^2 \times 10 \text{ cm}$ plastic scintillators	Fig. 1a
Alto [2]	La Paz (Bolivia) 4 200 m.a.s.l.	Same as above	Fig. 1b
Volcano Ranch I [3]	Albuquerque (N.M.) 1 800 m.a.s.l.	$4 \cdot 0.9 \text{ m}^2 \times 10 \text{ cm}$ plastic scintillators	Fig. 1c
Volcano Ranch II	Same as above		Fig. 1d

are enough to reconstruct the shower size within a factor better than two, provided that the number of particles detected at each point is larger than at least 20. The condition of 100% efficiency, computed for vertical showers for the minimum size N to be detected, fixes the area and the relative spacing of the detectors. The total number of detectors, and so the area of acceptance, is related to the minimum counting rate necessary to reach a significant statistics in a reasonable time.

A general condition on the type of detector used for the «total density sampling» is to be such as to not produce transition effects for the particles in the shower front. For this reason, light element detectors such as plastic scintillators have been used.

In Table I are listed the main data concerning the M.I.T. «E.A.S.» experiments for the investigation of the primary Cosmic Ray energy spectrum, and in Fig. 1 sketches are given

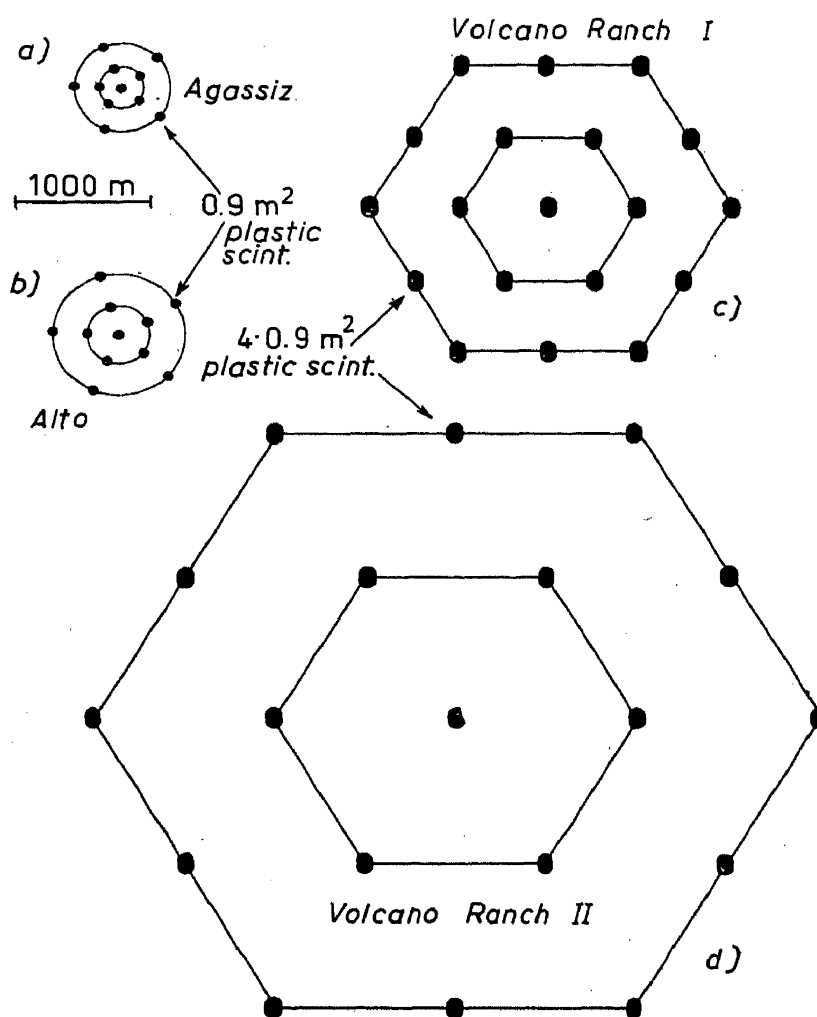


Fig. 1. — E.A.S. arrays.

Max. area	Range investigated	Time « on »	Nb. of showers accepted for analysis
0.2 km ²	$5 \cdot 10^{15} \text{ eV} \div 5 \cdot 10^{18} \text{ eV}$	20 months	500
0.35 km ²	Same as above	Still in operation	
2.5 km ²	$5 \cdot 10^{17} \text{ eV} \div 10^{19} \text{ eV}$	6 months	700
10 km ²	$> 10^{19} \text{ eV}$	Still in operation	

of the geometrical disposition of the different arrays. Figure 2 illustrates the block diagram of the electronics associated to the array Volcano Ranch I; a similar set-up is used in the other experiments. The shower data, derived

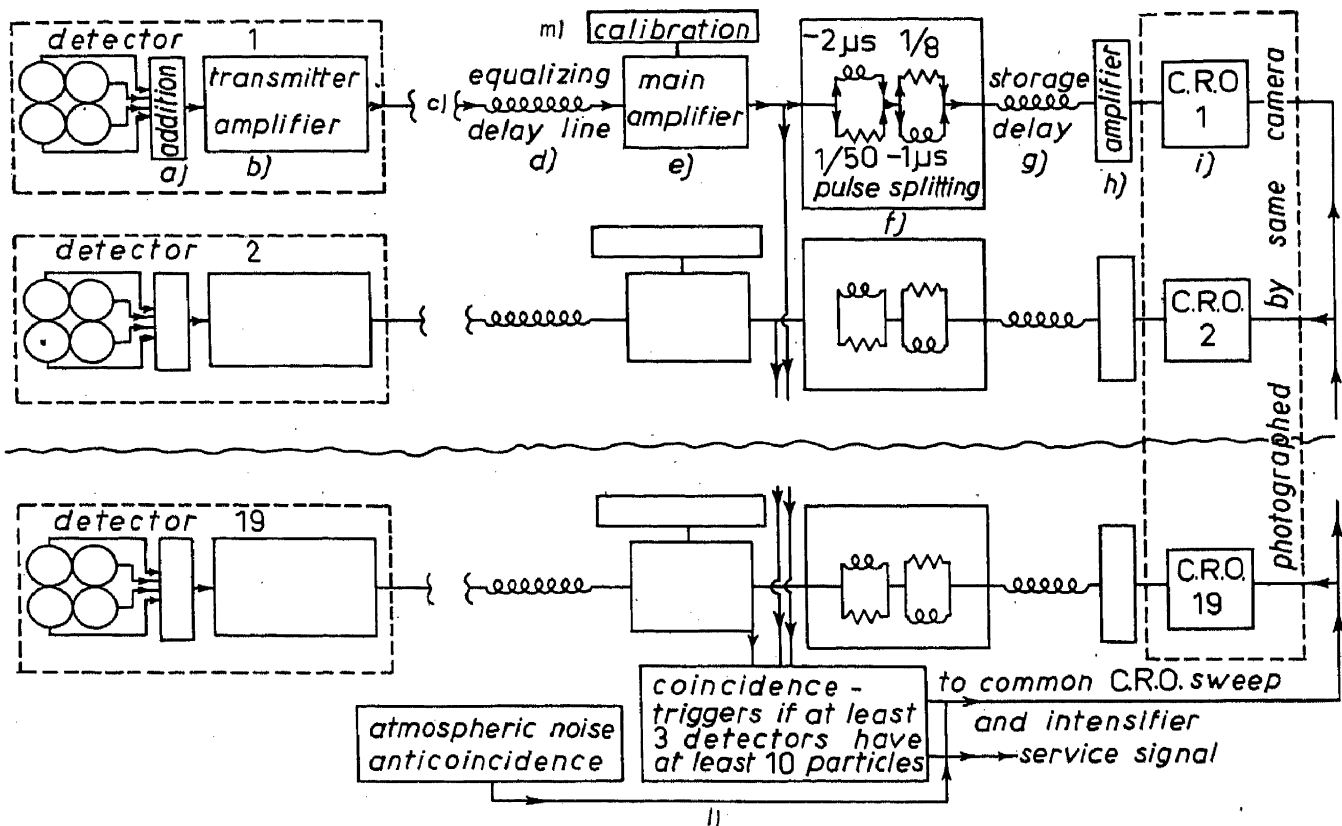


Fig. 2. - Volcano Ranch I Electronics Block Diagram. The signals from the four sections of each detector are added (a) and transmitted by coaxial cable (b, c); the electrical length of the cables from the different detectors are equalized (d); after amplification (e), the pulse goes to a pulse splitting circuit (f) which allows a dynamical range 1 ÷ 5000 particles through the 4 channels: I) attenuation by 400, no delay; II) attenuation by 50, 1 μs delay III) attenuation by 8, 2 μs delay; IV) no attenuation, 3 μs delay. The signal is then stored for ~ 5 μs (g) and after a last stage of amplification (h) reaches the cathode ray oscillographs (i). The coincidence circuit (l) triggers the common C.R.O.'s sweep and intensifier. The detectors are continuously calibrated (m).

from the scanning of the film monitoring the Cathode Ray Oscillographs connected to the individual detectors, are processed by an high speed electronic computer.

3. - Recent results on E.A.S.

The results quoted in the following have to be considered as very preliminary, the complete analysis not having yet been completed; they refer to a six months run of the array Volcano Ranch I [3]. Figure 3 gives the Size differential vertical spectrum. The integral energy spectrum given in Fig. 4 is deduced from a cascade model elaborated by S. OLBERT [4].

An immediate feature of the spectrum is the absence of any indication of a cut-off at least up to 10^{19} eV. The two largest showers observed have $N \sim 10^{10}$, which corresponds to a primary energy $> 10^{19}$ eV.

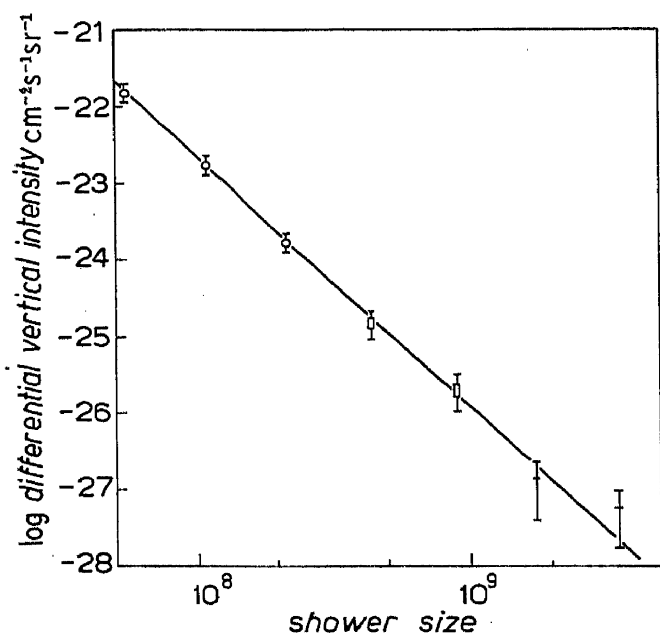


Fig. 3. — Shower size differential vertical spectrum (Volcano Ranch I).

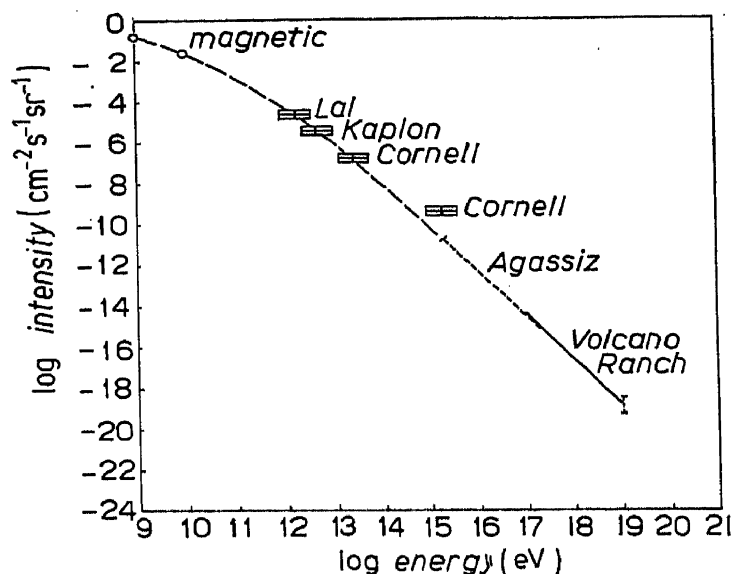


Fig. 4. — Primary integral energy spectrum.

These two facts combined seem to us a very strong evidence for the existence of an extragalactic component in the Cosmic Radiation. Extreme fluctuations can be reasonably excluded and, also if the primaries of all the large showers were heavy nuclei, their confinement in the galactic halo (not to speak of the galactic disk) should result by an unlikely combination of numerous factors.

As for as the arrival directions are concerned, primaries seem to be isotropic at least up to $5 \cdot 10^{18}$ eV, where the statistics is high enough to draw some definite conclusions.

REFERENCES

- [1] G. W. CLARK, J. EARL, W. L. KRAUSHAAR, J. LINSLEY, B. B. ROSSI, F. SHERB, D. W. SCOTT: *Phys. Rev.*, **122**, 637 (1961).
- [2] G. W. CLARK, I. ESCOBAR, J. HERSIL, S. OLBERT, D. W. SCOTT: *Proc. of the Intern. Conf on Cosmic Rays-Kyoto* (Sept. 1961). To be published.
- [3] J. LINSLEY, B. ROSSI and L. SCARSI: *Proc. of the Intern. Conf. on Cosmic Rays-Kyoto* (Sept. 1961). To be published.
- [4] B. ROSSI: *Proc. of Intern. Conf. on Cosmic Rays* (Moscow, 1959).

Electrons in the Galactic Cosmic Radiation.

B. DAYTON

University Institute for Theoretical Physics - Copenhagen

1. - Introduction.

It has long been recognized that at least some electrons should be found among the galactic cosmic rays. For example, electrons are produced, via π - μ -e decay, from the secondaries of collisions between energetic protons or nuclei and the interstellar gas. In such a collision about 5 % of the incident energy, on the average, is given to electrons (and positrons). However, it is known from the low relative abundance of the light nuclei (and from the existence of the very heavy nuclei) in the primary cosmic radiation that not more than a few grams cm^{-2} of the interstellar matter has been encountered by an average cosmic ray particle. Thus it would seem that not more than a few tenths of one percent of the galactic cosmic ray energy should be carried by electrons unless there are additional electron sources. Indeed, one might expect even less because of the additional energy losses to which high energy electrons are subject.

Until very recently the only relevant experimental evidence has been the null result [1] of CRITCHFIELD, NEY, and OLESKA, published in 1952. In their balloon-borne cloud chamber the number of observed events attributable to incident electrons and γ -rays was smaller than the number they expected as secondaries from interactions in the residual atmosphere. They concluded that not more than 0.6 % of the incoming particles having energy greater than 1 GeV are electrons or γ -rays.

The recent rapid advances [2] in radio astronomy have provided new evidence for the existence of highly relativistic electrons throughout the galaxy. It is now generally believed that the cosmic radio noise is due to the «synchrotron» radiation (also called magnetic bremsstrahlung) emitted by high energy electrons as they follow spiral paths about magnetic field lines in the galactic disc and halo. From the observed intensity of the radio noise in the 3 m band it has been estimated that electrons in the galaxy have an average

energy density of about 10^{-2} eV cm $^{-3}$ in the GeV range. This compares with about 1 eV cm $^{-3}$ for the energy density of the galactic cosmic rays measured at our particular place in the galaxy. If we assume that the sun does not lie in an inaccessible region and that the electrons (as well as the nuclear particles) may diffuse more or less freely between disc and halo, we might expect something like 1% of the non-solar particles reaching earth to be electrons. Such an assumption clearly places limitations on any model for the propagation of cosmic rays in our galaxy, so a measure of the electron fraction is of high interest.

Recently two measurements [3, 4] have been reported by EARL and by MEYER and VOGT, respectively. The first is essentially a repeat of the earlier experiment by CRITCHFIELD *et al.*, but with higher-flying balloons. The second is a counter experiment, also at balloon elevation. These experiments, which will be discussed in some detail in a following section, indicate that of the order $(1 \div 3)\%$ of the cosmic rays incident at the top of the atmosphere are electrons, though it is not clearly established that these are of galactic origin.

2. - The evidence from radioastronomy.

The stars and clouds of interstellar gas are largely confined to a disc about 10^5 l.y. (light years) in diameter and something less than 10^3 l.y. thick. The structure and motions of the spirals arms, along which the stars and the gas are concentrated, have been revealed principally through observations [5] of the 21 cm line of neutral hydrogen.

However, it is the radiation in the continuum with which we are concerned. Except for some discrete « thermal » sources, such as the planetary nebulae, the continuous radiation is characterized by a power-law frequency spectrum rather than by the Planck distribution expected for a radiating hot gas. This continuous, non-thermal radiation is emitted strongly from some discrete sources, moderately from the whole galactic disc, and weakly from a nearly spherical halo (polar radius ≈ 0.6 equatorial radii) surrounding the disc and having a volume of the order $V \approx 3 \cdot 10^{68}$ cm ^{3} . (Direct observations of the halo about our sister galactic system, M31 in Andromeda, made with the large radio telescope at Jodrell Bank, give strong support to the above description of our own galactic envelope.)

The boundary of the halo is probably diffuse but within the volume V continuous non-thermal radiation [6] is emitted more or less uniformly and with an intensity $(10 \div 20)\%$ of that from the region of the disc. The spectral distribution is given by the power law:

$$I(\nu) d\nu = k\nu^{-\alpha} d\nu,$$

where $\alpha = 0.7 \pm 0.1$, from both disc and halo, over a range of wave lengths from decimeters to 30 meters. The estimated power of this non-thermal radiation from the galaxy is $P \approx 10^{38}$ erg s⁻¹, more than 80 % of which comes from the halo.

SHKLOVSKY, in 1956, showed that the observed cosmic radio emission can be accounted for by assuming a galactic halo, filled with a tenuous highly ionized gas, whose concentration is of the order $(10^{-3} \div 10^{-2})$ cm⁻³, and into which large scale magnetic fields ($B \approx 2 \cdot 10^{-6}$ to $5 \cdot 10^{-6}$ Oe) are frozen. In this vast volume, electrons of an energy density $w \approx .01$ eV cm⁻³, and with individual energies in the GeV range, diffuse along helical paths, emitting synchrotron radiation. From the observed frequency spectrum it is possible to deduce that the electrons have a power law energy spectrum similar to that of the galactic cosmic rays observed at earth. (The relevant relationships between magnetic field, electron energy, and frequency of radiation are outlined in the next section.)

Postulating such a magnetic field-filled halo is just what is needed as a containment volume for cosmic rays and makes possible a theory of the galactic origin of the cosmic rays. (Whether or not the highest energy cosmic rays can be contained within the galactic halo depends upon the dimensions and field strengths of the largest more-or-less regular field regions).

The halo medium has not been directly observed, but some properties for it are suggested by astrophysical considerations. In order for the halo to be held out in an apparently stable shape against the gravitational attraction of the mass concentrated in the disc, it is necessary to ascribe to the gas particles mean random velocities of the order 100 km s⁻¹. This corresponds to a kinetic temperature of the order $T \approx 10^6$ °K, so the gas is almost completely ionized. Such a plasma of (mainly) protons and electrons has such a high conductivity that, with the large linear dimensions available, even a rather turbulent magnetic field configuration would have a decay time long compared with the age of the galaxy.

If one assumes equipartition between the kinetic energy of the plasma particles and the magnetic field energy:

$$\frac{1}{2} \rho v^2 = \frac{B^2}{8\pi},$$

from which

$$\bar{B} \approx (2 \div 5) \cdot 10^{-6} \text{ Oe.}$$

The larger value, corresponding to an ion concentration of 0.1 cm⁻³, gives a total magnetic field energy in the galaxy:

$$W_B \approx 3 \cdot 10^{56} \text{ erg.}$$

This is the same as the estimated total energy W_{cr} of the cosmic rays in the galaxy [7], assuming our observations at earth to provide a fair sample. This is stated only as an interesting observation as there is no obvious mechanism for achieving an equipartition between the cosmic rays, the magnetic field, and the plasma.

An independent hint that an ion concentration, $n = 10^{-2} \text{ cm}^{-3}$, is about right comes from the observed « break » in the emission spectrum at a frequency $\nu \approx 10^7 \text{ s}^{-1}$. The electrons dominantly responsible for emission at lower frequencies would lose energy by collisions with the plasma at a greater rate than by « synchrotron » radiation.

The most convincing evidence that the cosmic radio noise is due to this latter process comes from detailed studies [8] of one of the discrete, non-thermal sources in our galaxy, the Crab nebula. Not only is the frequency-spectrum given by a power law over more than 7 decades in frequency, but the radiation is strongly polarized, both in the visible and in the radio wave lengths. The direction of polarization is rather regular except that it is everywhere perpendicular to the envelope at the boundaries of the nebula, just as would be expected for the « synchrotron » radiation of electrons with energies up to 10^{11} or 10^{12} eV in a magnetically closed region.

The mechanism of the electron acceleration process is not clear, but it is likely that it would affect protons and heavier nuclei as well, so it is assumed that there are also energetic nuclear particles in the same region. In the latter case, because of their much larger masses, the « synchrotron » radiation is negligible. The special role assigned to such sources as the Crab nebula in the so-far most successful approach to a theory of the origin of the cosmic rays [7, 9-12] will be discussed later in connection with the sources of the galactic electrons.

3. - The theory of synchrotron radiation:

The fact that an accelerated charge radiates follows directly from classical electromagnetic theory. SCHWINGER [13] carried out the calculation for the case of an electron moving in the uniform magnetic field of a synchrotron; hence the the name. The application to our problem was carried out by GINZBURG [9]. Here we only assemble the relevant formulas, all quantities being expressed in C.G.S. electrostatic units unless otherwise specified.

The angular frequency of an electron of energy $E \gg mc^2$ in a magnetic field is

$$\omega = \frac{eB_{\perp}}{mc} \left(\frac{mc^2}{E} \right) \text{ s}^{-1},$$

where $B_{\perp}$ is the component of the magnetic field perpendicular to the velocity of the electron. (In the halo a 1 GeV electron would spiral with a radius of the order one-tenth of an astronomical unit.)

The energy radiated per second in the frequency range ν to $\nu + d\nu$ is

$$P(\nu, E) d\nu = AB_{\perp} f\left(\frac{\nu}{\nu_{\max}}\right),$$

where the function $f(\nu/\nu_{\max})$ is a peaked but asymmetrical function about $\nu = \nu_{\max}$, with the maximum intensity occurring at

$$\nu_{\max} = .07 \frac{e}{mc} B_{\perp} \left(\frac{E}{mc^2}\right)^2 \text{ s}^{-1}.$$

The total rate of energy loss is obtained by integrating over all frequencies:

$$-\frac{dE}{dt} = \int P(\nu, E) d\nu = \frac{2}{3} c \left(\frac{e^2}{mc^2}\right)^2 B_{\perp}^2 \left(\frac{E}{mc^2}\right)^2 \text{ erg s}^{-1}.$$

From this one obtains the time required for an electron to radiate away one-half of its energy; expressed in convenient units,

$$T_{\frac{1}{2}} = \frac{.008}{EB_{\perp}^2} \text{ year},$$

where E is in GeV and $B_{\perp}$ in Oe.

One thus computes the half life of the $(1 \div 3)$ GeV electrons in the galactic halo to be in the range $(10^8 \div 2 \cdot 10^9)$ years, for the range $B = (2 \div 5) \cdot 10^{-6}$ Oe, taking only «synchrotron» radiation into account.

It can be shown that electrons, distributed in energy with a power-law spectrum,

$$N(E) dE = K E^{-\gamma} dE,$$

produce synchrotron radiation distributed in frequency with another power-law spectrum,

$$I(\nu) d\nu = k B^{\alpha+1} \nu^{-\alpha} d\nu,$$

where $\alpha = (\gamma - 1)/2$, so that the spectrum of electrons responsible for the cosmic radio noise may be obtained directly from the observed frequency spectrum.

4. – Sources of the electrons and their energy losses.

The «synchrotron» radiation from the Crab nebula is characterized by $\alpha = 0.35$, so the electron energy is distributed according to $\gamma = 1.7$. From the approximately known distance to the Crab nebula $((4\,000 \div 6\,000) \text{ l.y.})$ and the measured intensity of radiation, its total rate of radiation is about: $-dW/dt \approx 5 \cdot 10^{36} \text{ erg s}^{-1}$. If it has been radiating at least that strongly during the 900 years since it flared into being as a super nova, the total energy given to its electrons could not be less than 10^{47} erg . Actually it must be $(5 \div 100)$ times larger than this number for the average magnetic field to be within its permissible limits of $(10^{-3} \div 10^{-4}) \text{ Oe}$. If such super novae occur only about one per 100 years in the galaxy, they can contribute a total electron energy of $(2 \div 30) \cdot 10^{48} \text{ erg s}^{-1}$.

When the magnetic field of the over-extended envelope finally merges into that of the surrounding region, these high energy electrons will be released with some fraction of their original energy. The estimated power radiated by the entire galaxy ($10^{38} \text{ erg s}^{-1}$) should be supplied by these electrons.

From energy considerations alone one may say that at least the right order of magnitude of electron energy is generated within the supernova envelopes. How they are accelerated is unknown, though one might expect statistical acceleration (some kind of Fermi mechanism) if they are injected with $E \gg mc^2$. Radioactive decay of the neutron-rich isotopes that are presumably present in vast quantities [14] at an early stage in the life of the super nova could provide injection.

The relative importance of the various modes of energy loss depends upon electron energy, magnetic field strength, and the assumed gas or plasma concentration. The modes are:

- a) collisions with atoms or with plasma electrons,
- b) collision bremsstrahlung,
- c) inverse Compton scattering, and
- d) «synchrotron» radiation.

Process d) is probably dominant in the range:

$$5 \cdot 10^8 < E < 10^{10} \text{ eV}.$$

The energy spectrum of electrons becomes steeper as they lose energy such that in the above energy range: $\gamma \rightarrow \gamma + 1$. Thus one might expect the electrons now trapped in the crab nebula with a spectrum of $\gamma = 1.7$ to eventually approach an energy spectrum with $\gamma = 2.7$ in the galactic halo. The radia-

tion from the halo, having a frequency spectrum $\alpha = 0.7$, is due to electrons that now have an energy spectrum: $\gamma = 2\alpha + 1 = 2.4$. This may be only a fortuitous agreement since the Crab is believed to be atypical in its frequency spectrum.

It is possible that the electrons responsible for the cosmic radio noise are mainly of secondary origin. In a nuclear collision between two protons about 30 % of the energy is given to pions, of which $\frac{2}{3}$ are charged. In the process of π - μ - e decay, about $\frac{1}{4}$ of the pion energy is retained by the final electron or positron, the rest being lost to neutrinos. Thus, approximately 5 % of the proton energy is transferred to the electron component with each collision. With heavier nuclei, the fraction is not much different.

The energy required by electrons to provide the estimated 10^{38} erg s $^{-1}$ of galactic «synchrotron» radiation could evidently be met by such a source; *i.e.* 5 % of $(10^{39} \div 10^{40})$ erg s $^{-1}$, the estimated rate that energy is lost by the nuclear component of the cosmic rays through collisions.

However, the problem remains that an average cosmic ray particle has not encountered more than a small fraction of a mean free path of matter during its life. Also it is expected that the electrons secondary to the nuclear component would have a spectrum not much different than that of the nuclei; but the subsequent travels of the electrons should steepen the spectrum markedly.

If the electrons in the galactic halo, and presumably those that might be observed at earth, originate in a primary acceleration process, either as electrons arising from ionization or from radioactive decay, they should be predominantly negative. If, on the other hand, they are secondary to a nuclear component within an accelerating nebula, in the gas of the spiral arms, or in the tenuous halo, the resulting electrons should be of both signs, with some positive excess due to the charge of the primary nuclei.

An experimental determination of the relative numbers of positive and negative electrons in the galactic cosmic radiation is thus urgently needed.

An experimental measurement of the energy spectrum of the galactic electrons in the GeV region should give a strong clue as to their history: whether or not they have a common source with the electrons responsible for the radio noise.

If the electrons originate in the disc and then diffuse out into the halo, passing through a transition in magnetic field strength from B_d to B_h , the electrons will become more nearly parallel to the direction of the field, assuming the adiabatic invariant $p^2 \sin^2 \theta / B$ is preserved, where θ is the angle between the electron's velocity and the magnetic field direction. Thus, since $B_{\perp} = B \sin \theta$,

$$B_{d\perp}^2 / B_d^3 = B_{h\perp}^2 / B_h^3.$$

But the intensity: $I(\nu) \propto B_{\perp}^{\alpha+1} = B^{1.7}$, so $I_h/I_a \approx (B_h/B_a)^{2.5}$, if the electron density were the same. But the steeper spiral in the lower field region implies a lower electron density, so the dependence of the intensity on the relative field strength would be even stronger. We might expect a field ratio of 1:2 to result in an emission ratio as small as 1:10.

An experimental measure of the absolute value of the flux of electrons density ($E > 1$ GeV) in the disc (*i.e.* near earth) may help to clarify this question.

5. - Recent measurements of cosmic ray electrons.

In the experiment of Earl, a multiplate cloud chamber was flown to within 4.5 g of the top of the atmosphere where random expansions were made. A total of 11 soft showers identified as initiated by electrons and 6 identified as initiated by γ -rays were obtained while 380 penetrating particles were observed. Assuming all of the γ -rays are secondary, he estimates that not more than 2 of the electron events are secondary.

With these very limited statistics Earl obtains a ratio of electrons to protons:

$$I_e/I_p = (3 \pm 1)\% \quad \text{for} \quad E > 0.5 \text{ GeV}.$$

The observed events had electron energies from 0.5 to 3 GeV as estimated by the size of the showers. The incident electrons all multiplied in the first or second plate [14], so it is clear that these events were mainly not of nuclear origin.

The author concludes from the observed intensity of electrons that they must come from other sources as well as from the secondaries of nuclear interactions between primary nuclear particles and the interstellar gas.

The method of the above experiment has the virtue that a flux of electrons is compared with a flux of nuclear particles for the same detector geometry; however, the energies of the nuclear particles are not measured, so it is not demonstrated that the electron and nuclear component are for the same magnetic rigidity. More than half of the electrons are in the interval 0.5 to 1 GV, a region of magnetic rigidity known to have a paucity of α -particles, hence only marginally describable as galactic.

Because of the small number of events one is hardly able to draw far-reaching conclusions, but as a provisional result it suggests that of the order of $(1 \div 3)\%$ of the incoming particles are electrons (positive or negative).

The random expansion of a cloud chamber is not a method well suited to obtain good statistics, because of the low duty cycle; but each event has a high information content.

The experiment [4] of MEYER and VOGT has the virtue of many counts, but the interpretation is by no means so clear-cut. The counter arrangement selects minimum-ionizing incident particles and measures the overall range of such particles and their charged secondaries up to a maximum depth of 100 g/cm² in Pb. The range measurement should detect electrons between a few MeV and about 1 GeV, but incident protons undergoing nuclear interactions, particularly of the charge-exchange type, here make a substantial background. It is not clear that events attributed to electrons of energy greater than about 400 MeV cannot equally well be accounted for as nuclear interaction events.

Below about 400 MeV the authors find a spectrum steeply rising towards the low energy end. The flights took place in the far North where the geomagnetic cut-off is very low but poorly determined. By adopting a suitable low energy cut-off something like 1% of the incident particles may be taken as electrons, but the cut-off remains somewhat arbitrary.

The instrument is not designed to give any information over about 1 GeV, and in fact it does not seem possible to rely on the upper half of that energy interval. About the low energy particles it is just not possible to make any clear identification as galactic electrons. It is well known that the particles in this range of magnetic rigidity have more to do with the sun than with the galaxy.

It is interesting to note [15] that on three separate flights very similar electron fluxes and spectra were observed, even though, on one of these occasions, the low energy proton flux was enhanced by a factor two as a result of a solar flare.

6. - Requirements of a critical experiment.

An experiment or group of experiments is needed to measure:

- a) the flux of primary galactic electrons,
- b) their energy spectrum (particularly above 1 GeV),
- c) the relative number of positive and negative electrons.

To insure that they are truly galactic one should require that:

- 1) the magnetic rigidity is greater than 1 GV,
- 2) the electrons should arrive at those latitudes and from those directions where the geomagnetic cut-off rigidity is less than the electron energy and not elsewhere. Since it is well known to be difficult to calculate these geomagnetic effects accurately, they could be determined experimentally by simul-

taneously measuring the intensity of α -particles with a detector sensitive over the same solid angle as the electron detector. This would provide continuous data for an absolute calibration of the electron intensity as well as the E-W asymmetry to be expected for each sign of electron charge. Alpha particles are preferable to protons for monitoring because of the absence of albedo.

3) The electrons should be measured at a time when all singly charged particles of similar magnetic rigidity are constant in time. The period of the next solar «minimum» would be a particularly suitable time for such measurements.

Probably much may still be learned from experiments done at balloon altitudes, flying at more than one latitude and using the E-W asymmetry as effectively as possible, but a critical experiment necessarily will be a satellite experiment.

7. - A proposed electron detector.

A suitable apparatus might consist of the following:

a) First the incoming particle passes through a gas Čerenkov counter, operated at a relatively low gas pressure so as to register only if the particle is charged and has an energy such that, say,

$$E/mc^2 > 20.$$

Thus protons are heavily discriminated against by making use of the steepness of the energy spectrum to rule out most of the particle flux. Two photomultipliers must be used in fast coincidence to eliminate the registration of noise pulses.

b) Following this is an electron shower detector, consisting of a sandwich structure in which scintillation or solid Čerenkov counters are used to sample the development of an electron-initiated shower through the first few radiation lengths of a dense, high- Z absorber. By measuring the shower size near the maximum of its development (after 4 or 5 radiation lengths), the energy of the incident electron may be rather well estimated without the great weight of a total absorption detector. Also, such an open-ended counter has no upper energy limit. The sampling of the shower at more than one point makes possible still further discrimination against events that are not of pure electromagnetic origin.

Provisional estimates are that such an apparatus, suitable for balloon flights, could be made to detect one electron against a background of several thousand

protons or heavier particles in the energy range above 1 GeV. It would be essential to calibrate the shower detector and to determine the overall discrimination of the apparatus by measurements at all angles in beams of artificially accelerated particles. The 28 GeV proton beam at CERN and any of several electron beams now available are eminently suitable.

Eventually such a detector, or its descendant, could be reengineered for satellite operation. A polar orbit below the belts of trapped radiation would make possible a complete study of all of the objectives *a*), *b*), and *c*).

REFERENCES

- [1] C. L. CRITCHFIELD, E. P. NEY and S. OLESKA: *Phys. Rev.*, **85**, 461 (1952).
- [2] Cf., articles by BURBIDGE, by V. L. GINZBURG and by F. HOYLE: *Paris Symposium on Radio Astronomy* (BRACEWALL, ed.) (Stanford, 1959).
- [3] J. A. EARL: *Phys. Rev. Lett.*, **6**, 125 (1961).
- [4] P. MEYER and R. VOGT: *Phys. Rev. Lett.*, **6**, 193 (1961).
- [5] Cf., article by OORT, op. cit. in ref. [2].
- [6] Cf., article by MILLS, op. cit. in ref. [2].
- [7] V. L. GINZBURG and S. I. SYROVAT'SKIJ: *Soviet Physics Uspekhi*, **3**, 504 (1961).
- [8] L. WOLTJER: *Bull. Astron. Inst. Neth.*, **16**, 39 (1958).
- [9] V. L. GINZBURG: *The Origin of Cosmic Rays*, in *Progress in Elementary Particles and Cosmic Ray Physics*, vol. 4 (Amsterdam, 1958), p. 339.
- [10] B. PETERS: *Proceedings of the Varenna Summer School «E. Fermi»*, This volume p. 1.
- [11] V. L. GINZBURG: *Some aspects of theory of cosmic ray origin in Proceedings of the Moscow Cosmic Ray Conference*, vol. 3 (Moscow, 1960), p. 196; I. S. SHKLOVSKY: *Achievements in radioastronomy and Radioastronomical theory of the origin of cosmic rays in Proceedings of the Moscow Cosmic Ray Conference*, vol. 3 (Moscow, 1960), p. 205.
- [12] P. MORRISON: *The Origin of Cosmic Rays*, in *Handb. d. Phys.*, vol. 46/1, S. FLÜGGE, ed. (Berlin, 1961), p. 1.
- [13] J. SCHWINGER: *Phys. Rev.*, **75**, 1912 (1949).
- [14] J. A. EARL: Private communication.
- [15] P. MEYER: Private communication.

γ -Spectrography in a Rocket.

J. GOUPIL

Service d'Electronique Physique, Centre d'Etudes Nucléaires de Saclay - Saclay

Introduction.

This paper describes a device for measuring and transmitting the γ rays spectrum at the top of the atmosphere in the range on 100 keV to 2 MeV.

This assembly has been studied and realized in the Centre d'Etudes Nucléaires de Saclay in France.

The detector was placed in the four-stage rocket ANTARES. The expected altitude was 350 km and the flight duration approximately 8 minutes.

The detector is a photomultiplier with a 1 in. $\times$ 1½ in. sodium iodide crystal (Fig. 1).

The elimination of the pulses produced by charged particles is accomplished by means of a plastic scintillator which surrounds the first scintillator.

When a coincidence is produced on the two photomultipliers, the signal is eliminated.

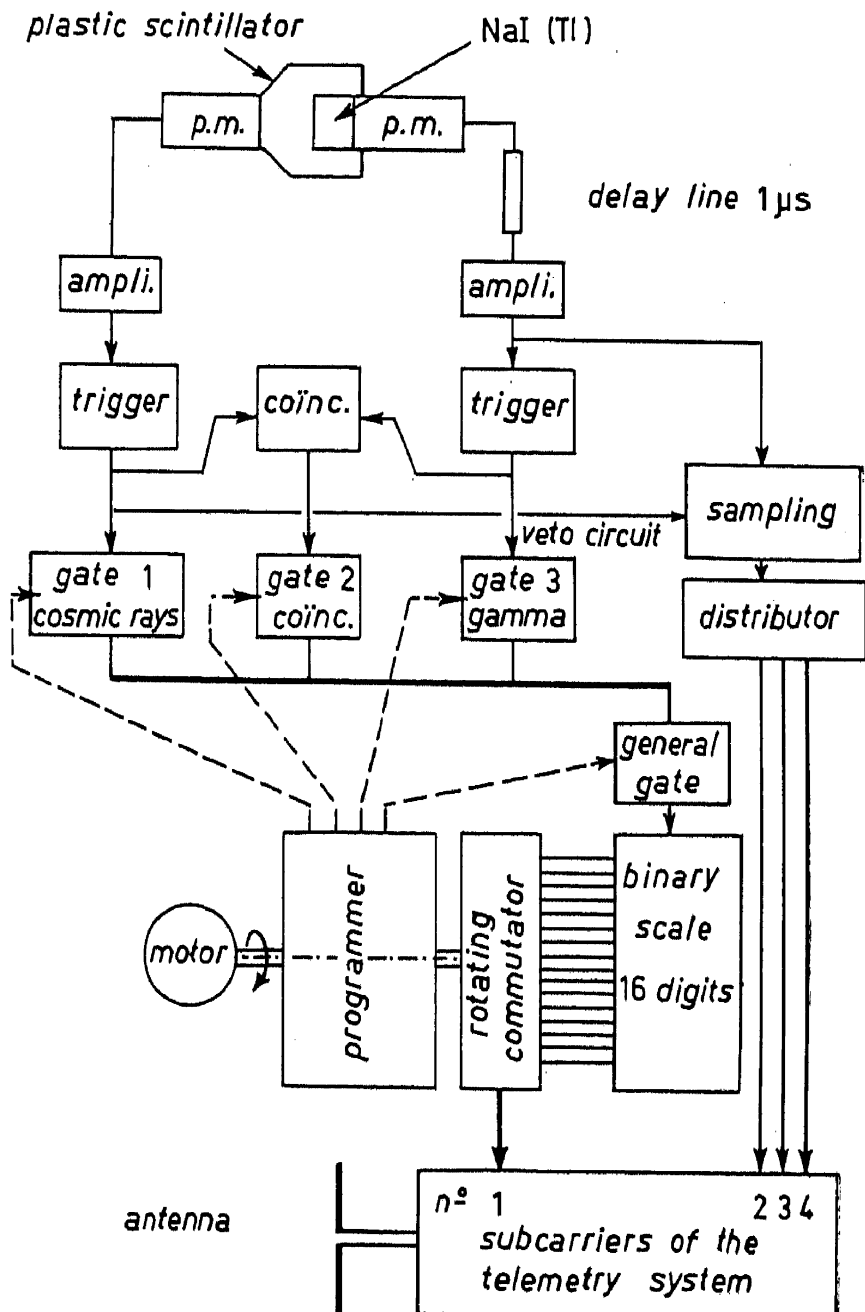


Fig. 1. - Block diagram.

We want to transmit to the ground:

- 1) a γ spectrum;
- 2) the counting rates of the two photomultipliers and the coincidences.

For this, the transmission device of the rocket is used. It consists of 5 frequency modulated sub-carriers. These sub-carriers modulate the amplitude of the high frequency wave to 100 MHz.

One sub-carrier is destined to the measurements of pressures, temperatures, accelerations and other parameters of the flight.

We have used one sub-carrier for counting and 3 sub-carriers for the transmission of the spectrum, because this measurement contains a large quantity of information.

1. - Counting rate.

The counts are sent successively on a binary scale containing 16 stages. The storage time is about one second, and the reading time one quarter second.

A mechanical commutator tests the state of the 16 stages successively. Then, after resetting to zero, the following counting is effected.

One can measure this way, the counting rates of one to about 65 000 counts per second.

2. - Transmission of the spectrum.

The band width of each sub-carrier is 400 Hz. In this way we cannot transmit more than 70 pulses-second if the amplitude is to be conserved with precision.

Therefore the electronic's must realize:

- 1) The sampling of the pulses leaving the γ photomultiplier.
- 2) The distribution of pulses on the three sub-carriers.
- 3) The increase of the pulse duration from 5 μ s to 10 ms.

Figure 2 shows how these operations are realized. The pulse is applied to a pulse-height to time converter through a linear gate.

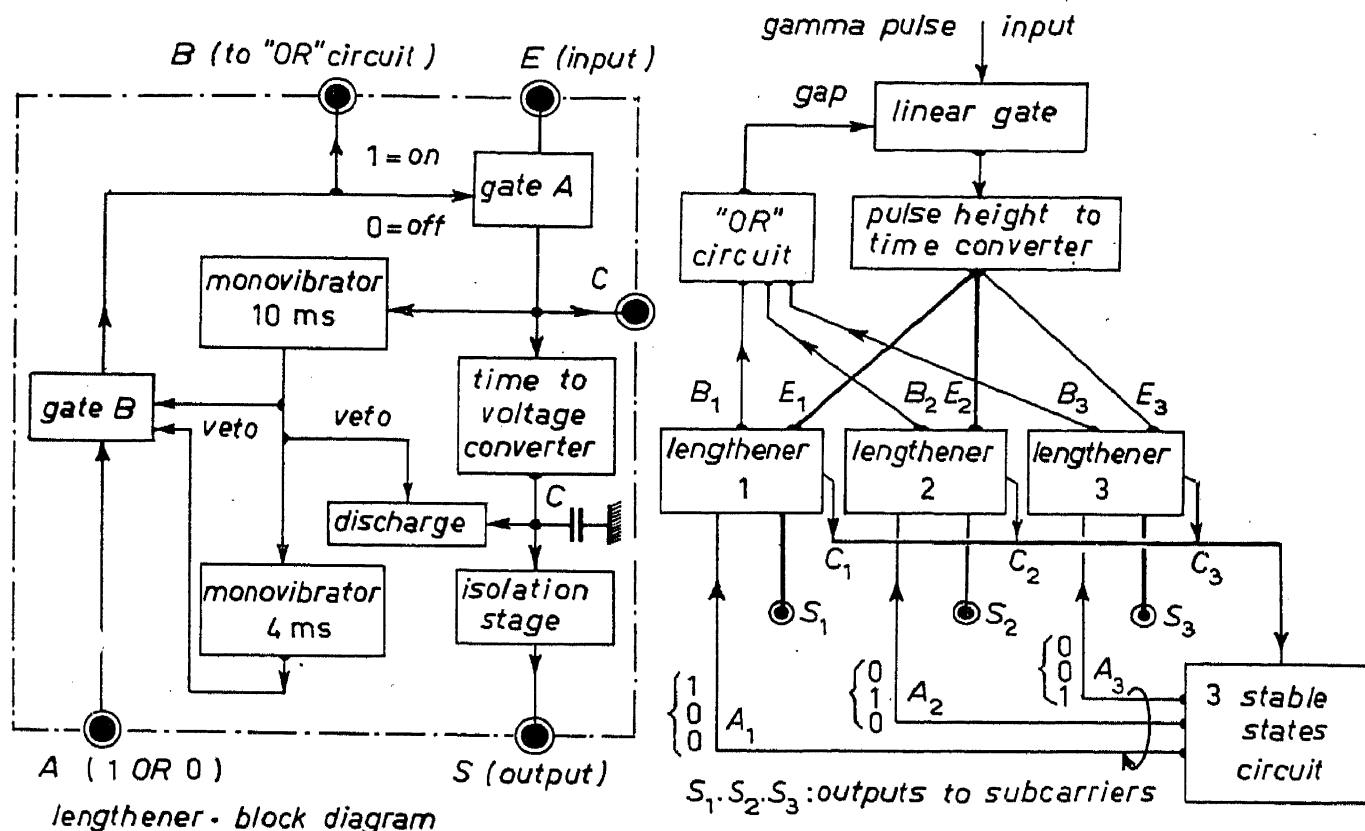


Fig. 2. - Sampling, distributor and lengthening.

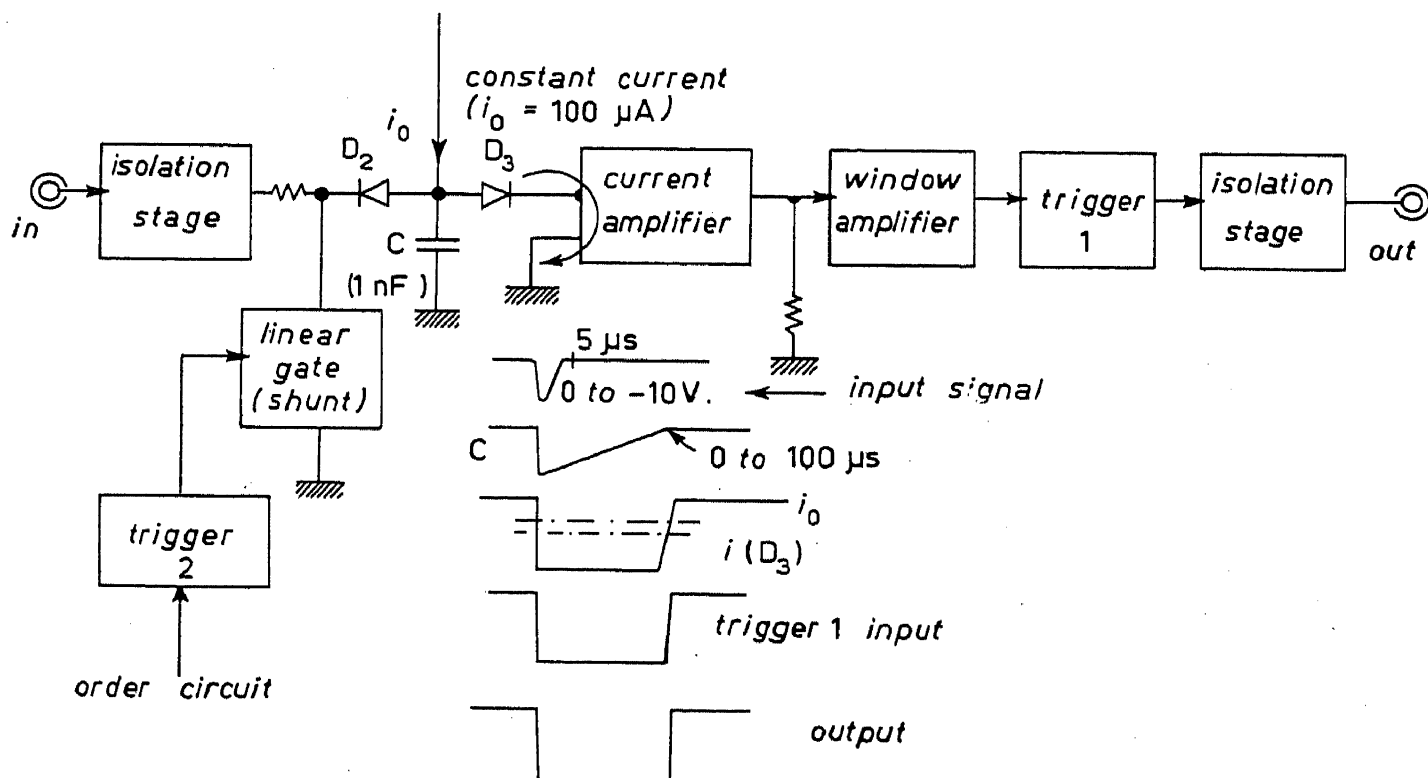


Fig. 3. - Pulse-height to time converter.

The output signal is then sent to 3 lengtheners. An electronic commutator permits the functioning of each lengthener separately.

If the lengthener is already in the process of treating a pulse, the linear gate is closed and no other pulse can reach it, as long as its work is not finished.

The lengthener consists of a gate, a time-amplitude converter and a capacitor used as a memory for 10 ms.

This time of 10 ms is given by a monovibrator. At the end of this time, the memory is reset to zero.

In addition a dead-time of 4 ms is provided between two pulses by means of a second monovibrator.

If one of these two monovibrators is functioning, any new pulse is refused by means of gate B.

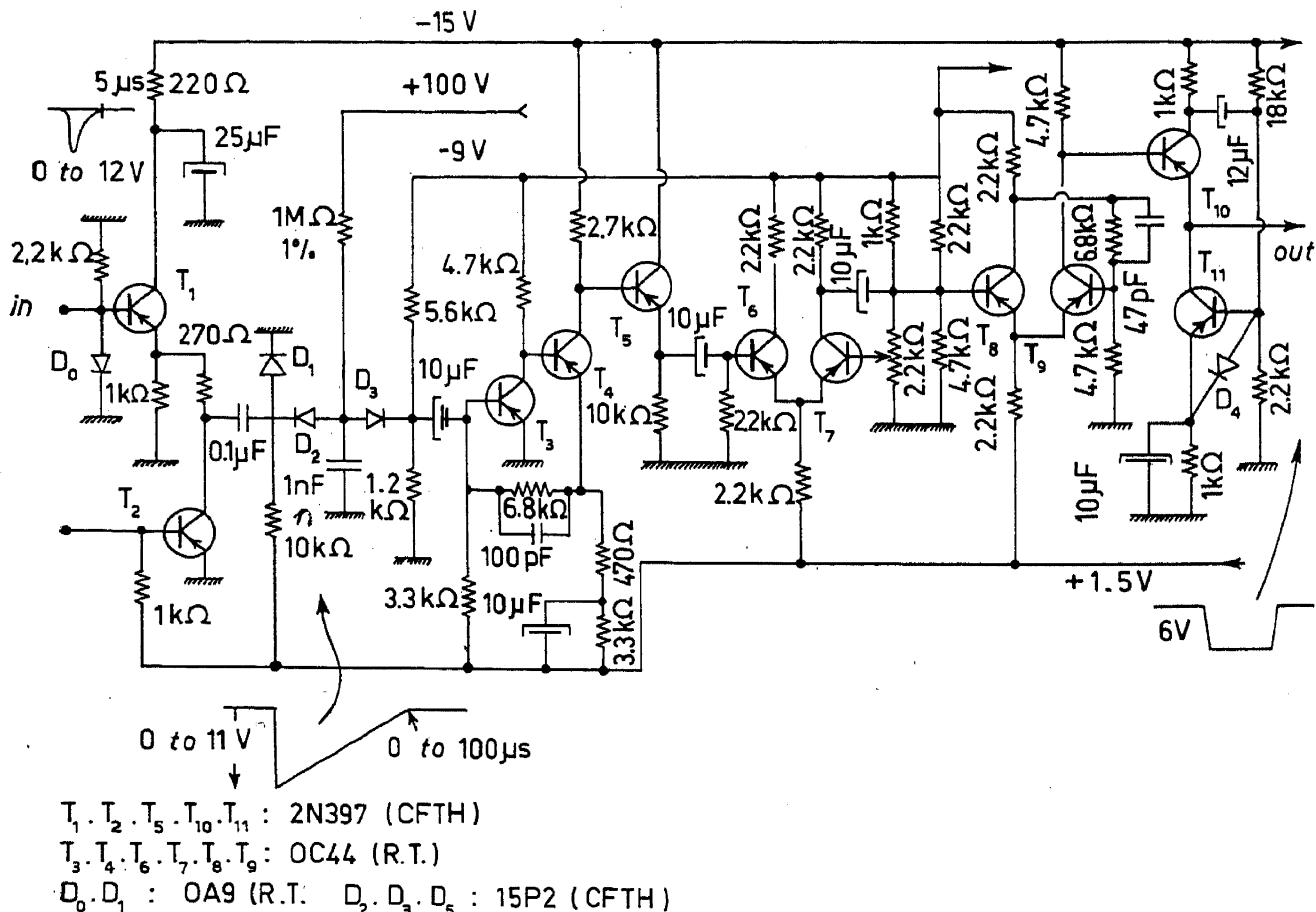


Fig. 4. - Pulse-height to time converter.

When a lengthener has received a pulse, it sends a pulse to the electronic commutator which autorizes the functioning of the following lengthener, (Fig. 2).

Figure 3 shows the principle of the pulse-height to time converter.

The pulse charges a condenser which is discharged by a constant current.

The time necessary for the point *C*, to come back to its initial state, is proportional to the amplitude of the input pulse.

For an amplitude of 10 V, there is a time of 100 μs.

Figure 5 shows the lengthener.

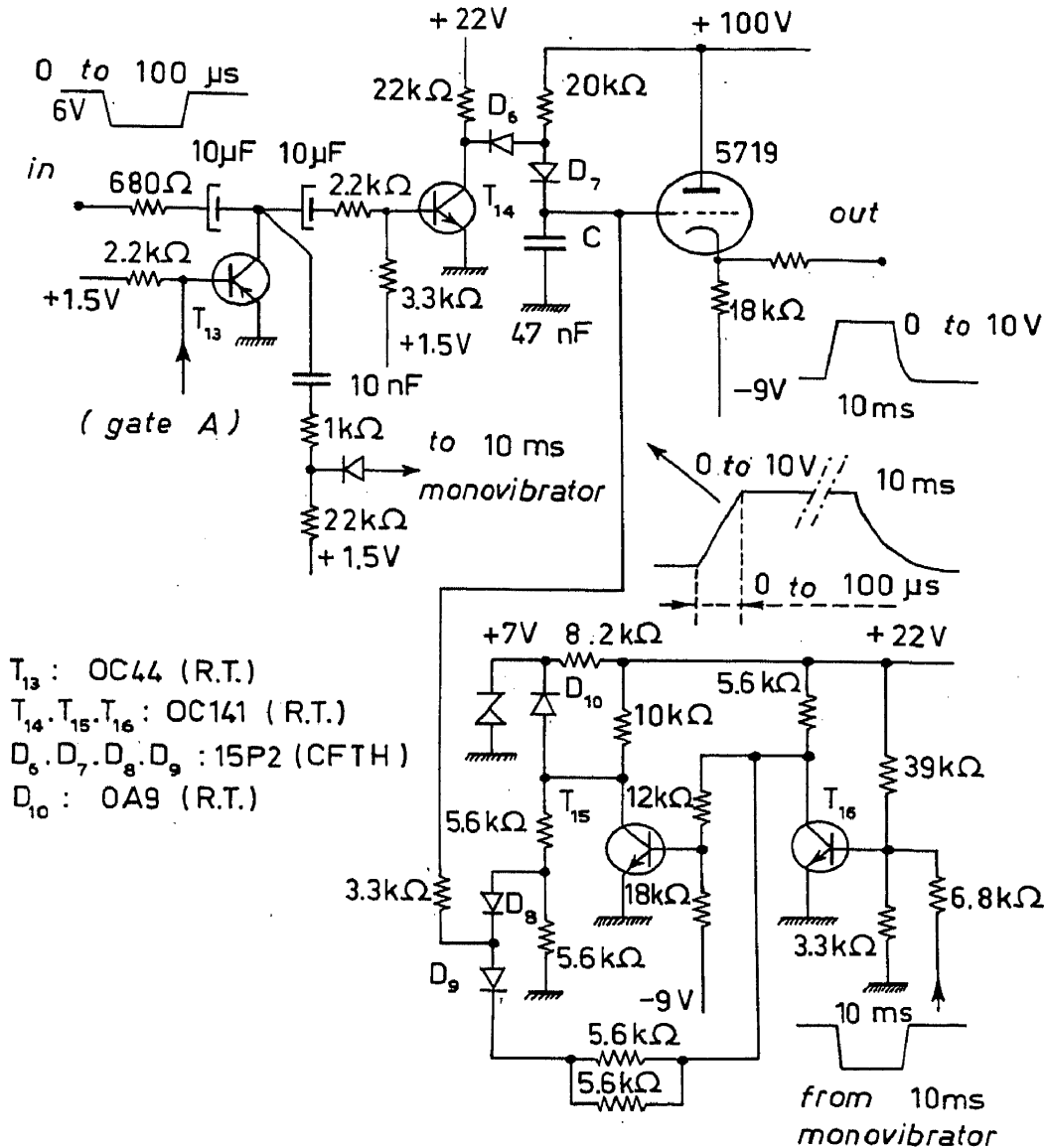


Fig. 5. - Lengthener.

C is the memory condenser. It is charged by a constant current of 5 mA during a time proportional to the amplitude of the pulse applied to the first converter.

The voltage obtained is therefore proportional to the pulse given by the gamma photomultiplier.

A cathode follower copies this voltage.

After 10 ms the condenser *C* is discharged by a very precise special device.

The linearity of the assembly: pulse-height to time converter and len-

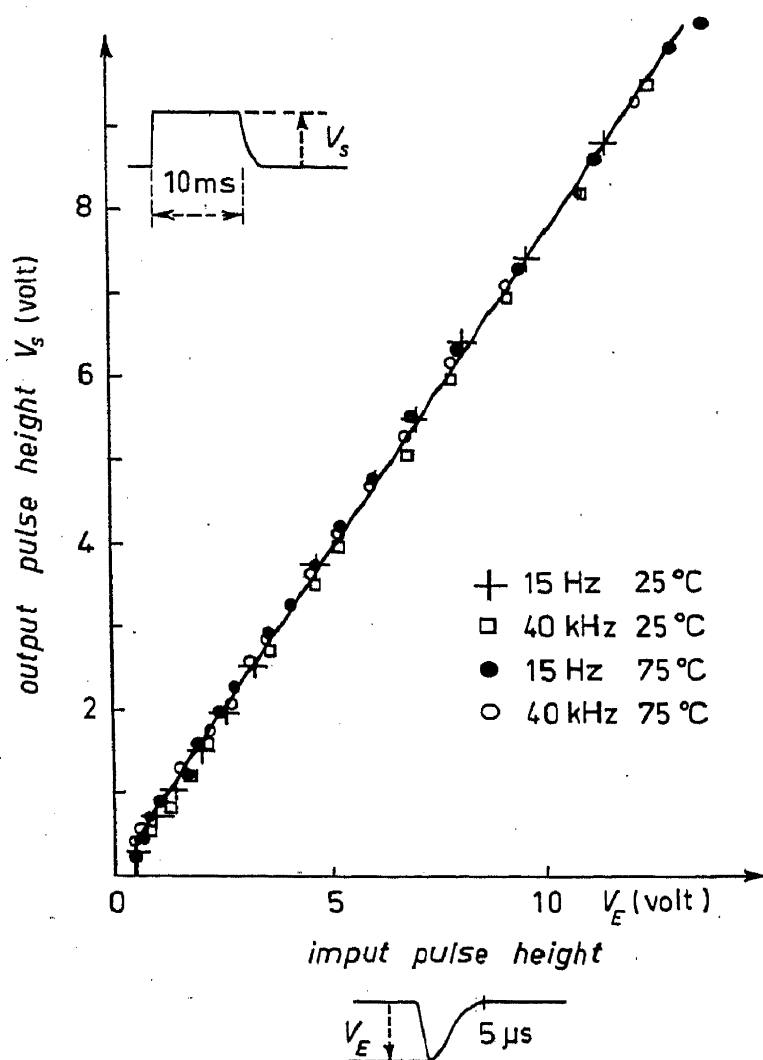
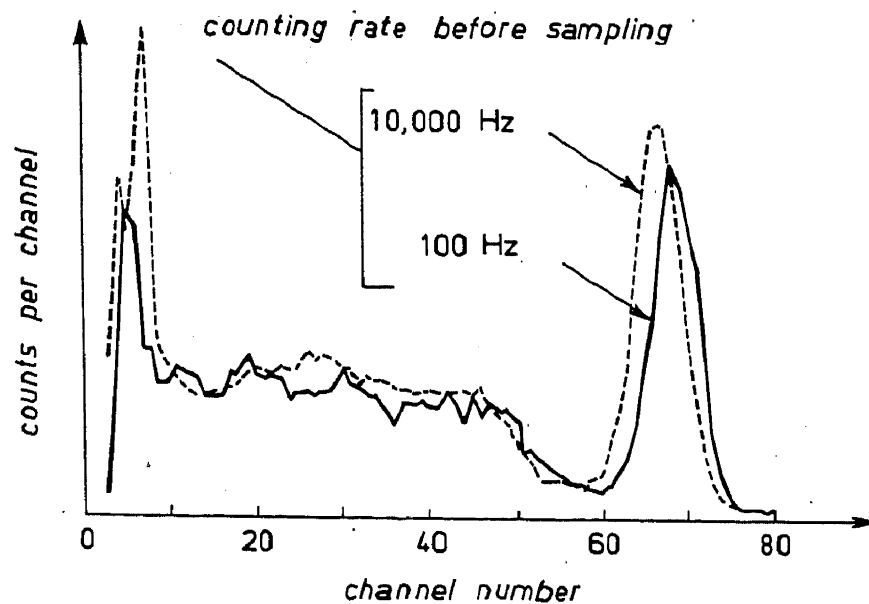


Fig. 6. — Linearity and stability.

theners, is given by Fig. 6 for two temperatures, 25 °C and 75 °C, and two pulse frequencies 15 and 40 000 counts per second.

Fig. 7. — ^{137}Cs spectrum after telemetry.

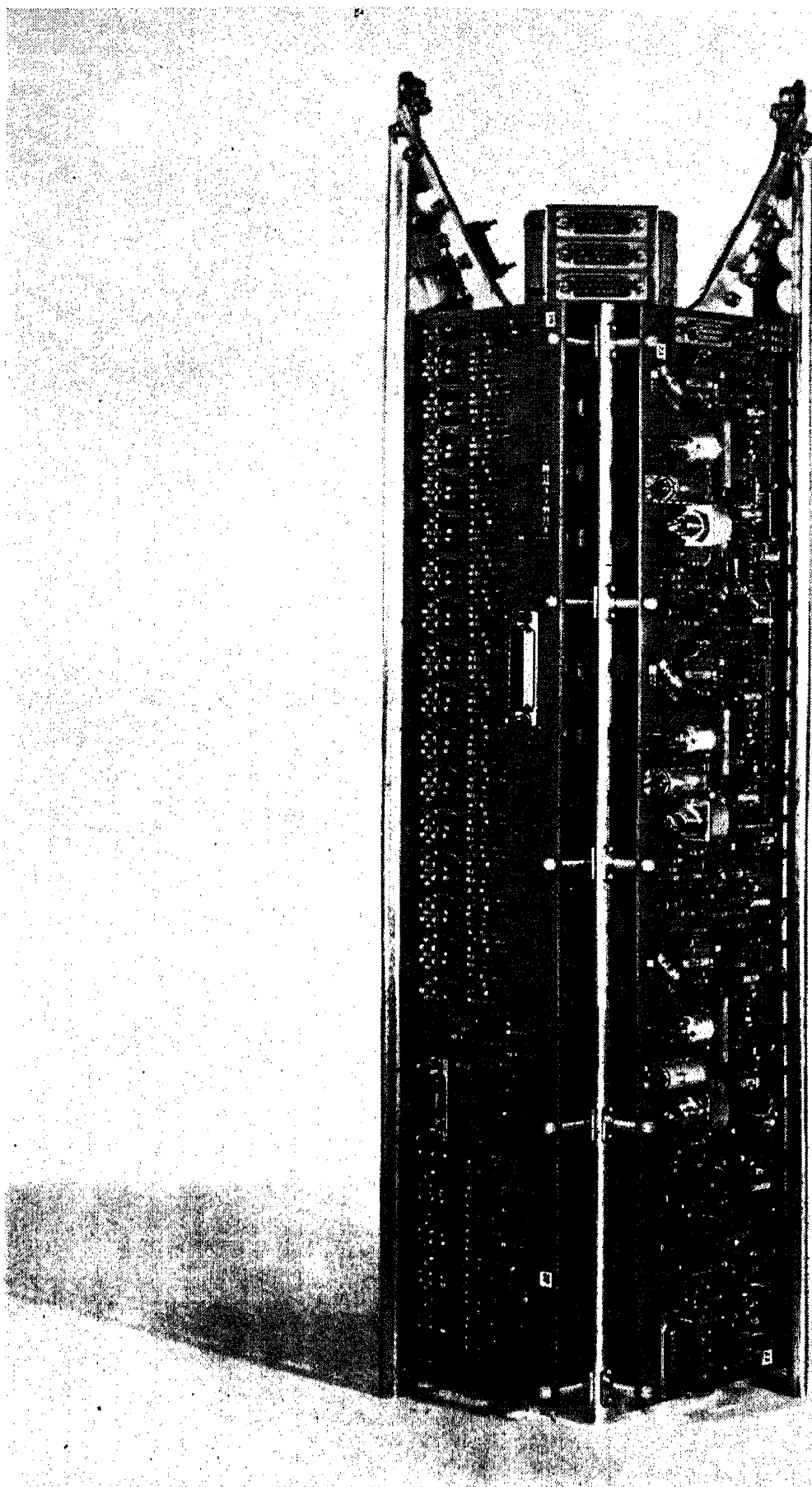


Fig. 8. — View of counting (on the left) and lengthening printed circuits.

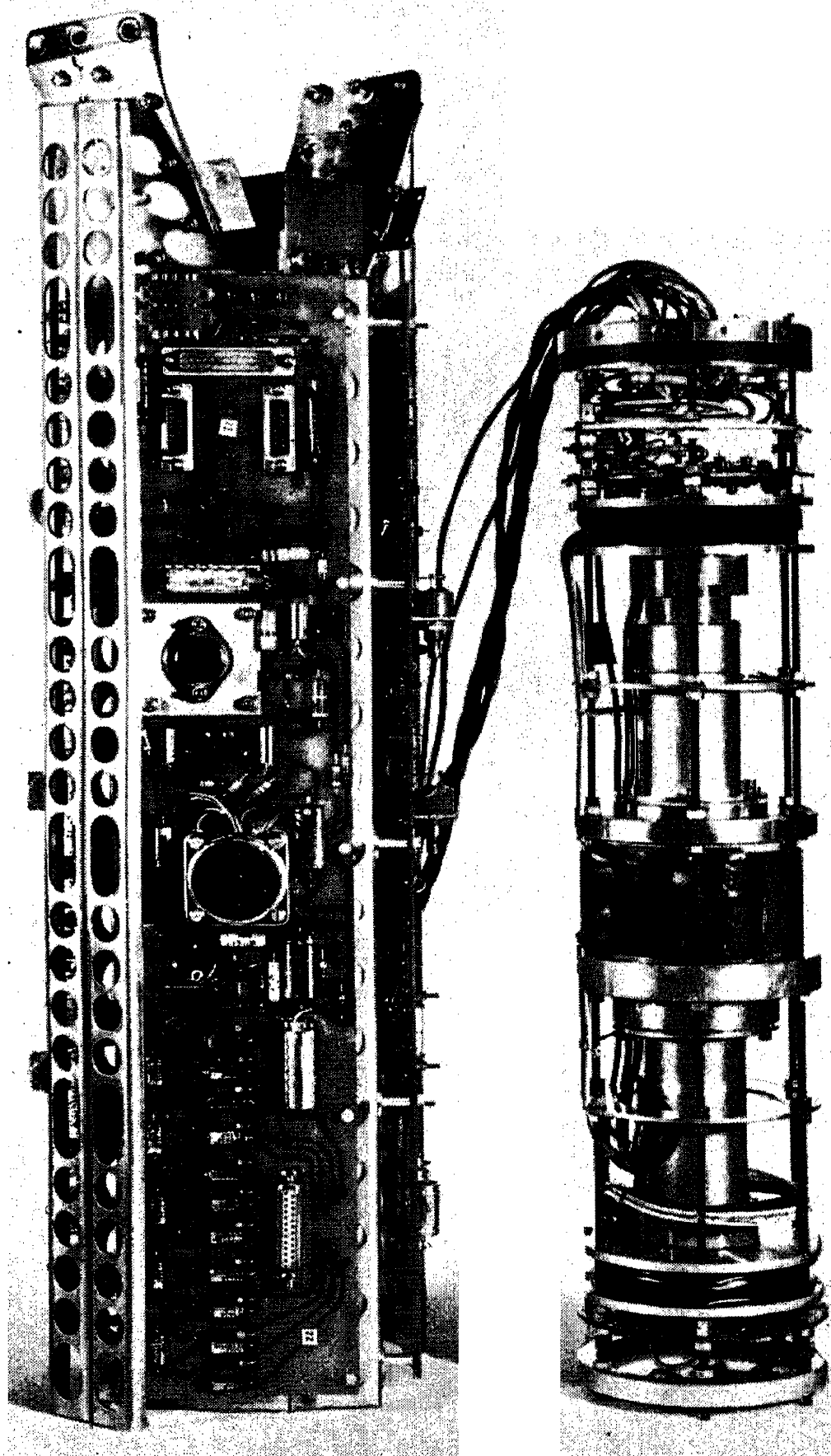


Fig. 9. - Mechanical device with regulated power supply and detector (on the right).

The spread of the points does not exceed 2%. Because of the sampling, the number of pulses transmitted by each sub-carrier cannot be larger than 70 per second.

Figure 7 shows a spectrum of ^{137}Cs transmitted by radio, for two counting rates: 100 and 10 000 counts per second. The spread of the peaks is 3%. Finally, Fig. 8 and 9 show two views of the assembly realized with printed circuits. There are 130 transistors, and 200 diodes.

The Interstellar Gas and the Galactic Halo.

H. C. VAN DE HULST

Leiden Observatory - Leiden

1. - Two new branches of astronomy.

The name « cosmic radiation », first used about 1909, was a double confession of ignorance. « Radiation » or « rays » meant that we did not know whether they were photons or fast particles. « Cosmic » meant that we did not know where they came from, but clearly the origin was somewhere outside the earth.

The term « cosmic static » was used from 1931 until it was replaced by the term « radio astronomy ». Again « cosmic » was a confession of ignorance and « static » meant radio noise.

Since those times there have been enormous advances in measuring techniques, in factual knowledge and in our understanding of these two phenomena. This rapid progress illustrates the rule that often in science an improvement in technique triggers off a whole chain of developments in experiment *and* in theory. Only those familiar with the unsolved problems usually can take up this challenge. As an example relevant to our subject I may mention the *Crab Nebula*.

This nebula had been known from spectral studies for a long time to consist of two parts: a shell of red filaments and a white nebulous centre. It had been vaguely thought that the white light from the centre was the radiation from a very hot gas but a consistent explanation could not be found. Radio measurements ruled this explanation out by showing that the flux density at radio wave lengths was about 100 times stronger than in the visual region. This led ŠKLOVSKIJ 1952 to take up suggestions made in a different context by ALFVÉN, HERLOFSON and KIEPENHEUER and to propose a completely different mechanism: the origin of the optical and the radio emission should be synchrotron radiation from relativistic electrons in a magnetic field. If the velocity of an electron is small compared with the velocity of light, it radiates only at the gyrofrequency but when the velocity is close to that of light it

radiates at a great number of harmonics and the spectrum may for practical purposes be regarded as continuous. ŠLOVSKIJ showed that the synchrotron process could account for both optical and radio observations of the *Crab Nebula*. He also suggested that it could be the origin of the continuous radio emission from the Galaxy and that it should be possible to detect linear polarization in this. Synchrotron radiation is essentially linearly polarized with the electrical vector perpendicular to the magnetic field. Soon after this suggestion, the white light from the central part of the *Crab Nebula* was indeed found to be strongly polarized (the best measurements have been made by WOLTJER on plates taken by BAADÉ) and very recently polarization has been found in the radio frequency radiation from the Galaxy.

It is interesting to note that the «cosmic static» has now become a recognized part of astronomy while «cosmic rays» are still regarded as part of physics. Both fields involved new techniques, new concepts and a new terminology which required a large effort from the ordinary astronomer. I suppose the difference is that he feels more at home with the final results of radio astronomy observations, which at least can be presented in the form of maps of the sky. But evidently cosmic rays should also be considered as part of astronomy, which—of course—is again part of physics.

The problem of getting acquainted with such new fields is comparable to that of learning a foreign language. In going to a new country it is absolutely necessary to speak and read *some* sentences in the other language. But when you feel well advanced, try the supreme test of hearing the insiders speak among themselves. Students from various fields are attending this course. I advise you to try this test and not to be content with things comparable to «the 100 most useful sentences» or «Reader's Digest». Go to the original publications and visit the research workers in their own institutes.

2. – Diagnostics of the interstellar gas.

My task here is to discuss what kind of information astrophysicists can give that is relevant to the study of cosmic rays and in this I shall limit myself to the conditions in space outside the solar system, *i.e.* away from the influence of solar «winds», «clouds» and their associated magnetic fields. The distances here are on a very much larger scale than in the solar system: instead of light minutes we have to deal with centuries.

The study of plasmas in the laboratory is now vigorously pursued in many countries in view of deriving power from controlled nuclear fusion. Many of these studies are useful to astronomy and have first been started in astronomy. The analysis and measurements of plasmas, now often called «plasma diagnostics», also very well describes the work of the astrophysicist when he tries

to find out something about the interstellar medium from its properties of emitting, absorbing and scattering light and radio waves. Of course, he cannot make new experiments but has to be content with what nature itself provides.

Both the time scale and the linear scale are « out of this world ». We observe structures of gas, stars, and magnetic fields, which are obviously very unstable, but changes are extremely slow. In the laboratory a plasma may change considerably in a few microseconds; out in galactic space we could take a picture every century and obtain a fascinating film.

The plasma physicists have discovered with regret that small impurities in the plasma have large effects on its behaviour. This phenomenon is well known to astrophysicists since 1926. Small amounts of oxygen ions (singly and doubly ionized) can be excited by collisions and return to the ground state by radiating visible light. This gives an effective sink of energy in the plasma. The radiation is usually in so-called forbidden lines, which can be produced only in very rarified gases where collisions are rare. At the same time intensity ratios between some of these emission lines are sensitive to the density and temperature of the plasma, and can be used for the « plasma diagnosis ». The singly ionized oxygen atom, for instance, emits a strong doublet in which the intensity ratio of the two emitted lines is a function of the electron density N_e .

$$\text{The ratio } \frac{I_{3729}}{I_{3726}} \begin{cases} = 0.47, & \text{for } N_e \geq 10^5 \text{ cm}^{-3} \\ = 1.5, & \text{for } N_e \leq 10^3 \text{ cm}^{-3} \end{cases}$$

The observed ratio can be used to determine N_e between these limits. Other ratios are more sensitive to electron temperature.

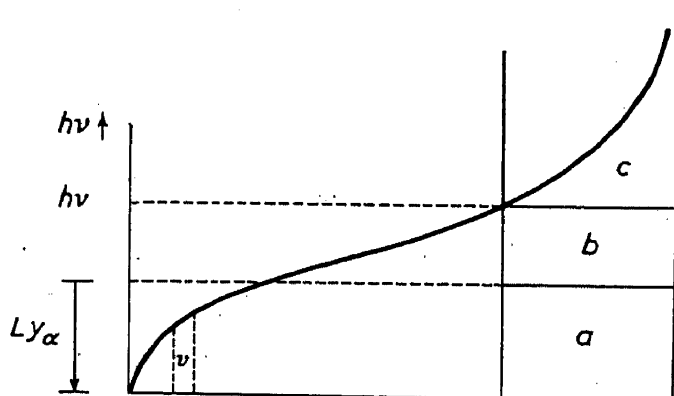


Fig. 1. — Diagram showing how ultra-violet energy of a hot star is converted and used up in the surrounding gas.

The energy budget of the gas surrounding a hot star is represented in a simple but unconventional way in Fig. 1. The usual way to draw the blackbody radiation curve is to plot $I(\nu)$ against ν so that the area under the curve represents total energy (in an appropriate unit). Figure 1 is derived by putting the quanta side by side like soldiers in the rank, each with its own height $h\nu$ and with unit width. In a formula, this means that the abscissa is

$\int_0^\nu (I(\nu')/h\nu') d\nu'$ and the ordinate $h\nu$, and the area under the curve again represents total energy.

The area under the curve to the right of the vertical line (*i.e.* regions $a+b+c$) represents the amount of energy available above the Lyman limit, *i.e.* in quanta which can ionize a hydrogen atom from the ground state. The excess energy $h\nu - h\nu_1$ initially appears as kinetic energy of one free electron. However, the electrons immediately redistribute this energy among themselves and the ions by elastic encounters and continuously transfer energy to the impurities like oxygen, which return to the ground state radiating in the forbidden lines. The electrons, which have lost most of their kinetic energy in this way, are finally captured by protons and return to a hydrogen atom in the ground state emitting Balmer lines and Lyman α . Or, if they combine directly to the ground state, the cycle is repeated. In equilibrium we thus may say that an amount of stellar radiation represented by the area a is converted into Lyman α , that of the area b in various other hydrogen lines and continua, including the Balmer series, and that of the area c is converted into radiation of the forbidden lines of the impurities.

Close to a hot star (type O or B) the intense radiation keeps the gas almost completely ionized, but further out this radiation gets exhausted and the gas is essentially neutral. The transition is very sharp and the gas surrounding the star is thus divided into an inner ionized region (« Strömgren sphere ») and an outer region where hydrogen is neutral. The brightness of the nebula surrounding the star (measured for instance in one of the oxygen lines) thus gives information about the area c ; direct observations of the central star give the area d (visible light). The ratio of these two is obviously a sensitive measure of the radiation temperature of the star.

These few examples may serve to show that our knowledge of the temperatures and densities of the interstellar gas is not based entirely upon theory but that a fair number of determinations directly from observation is possible. Doppler shifts and Doppler broadening, unfortunately, rarely give any useful information about the temperature because they are determined mainly by fluid motions and cloud velocities.

The first detection of the 21 cm line of H in 1951 has opened an enormous range of new observational possibilities to study the interstellar gas. The next important step will be to detect H_2^+ and H_2 from satellites.

3. – Dynamics.

Almost any problem in modern astrophysics leads to problems of dynamics and evolution. In most regions of the interstellar medium there may not even be a quasi-equilibrium which characterizes, for instance, the atmosphere of a star before the supply of energy gets exhausted. The implications for the problem of the origin of cosmic rays are numerous and although these impli-

cations are not dealt with in this paper, it may be good to review some of the facts as they are known from optical and radio studies.

The main part of the galaxy containing most of the stars is shaped like a disk with a diameter about 80 000 light years and with a thickness of less than 1/100 of this diameter (Fig. 2). The system rotates not like a solid body

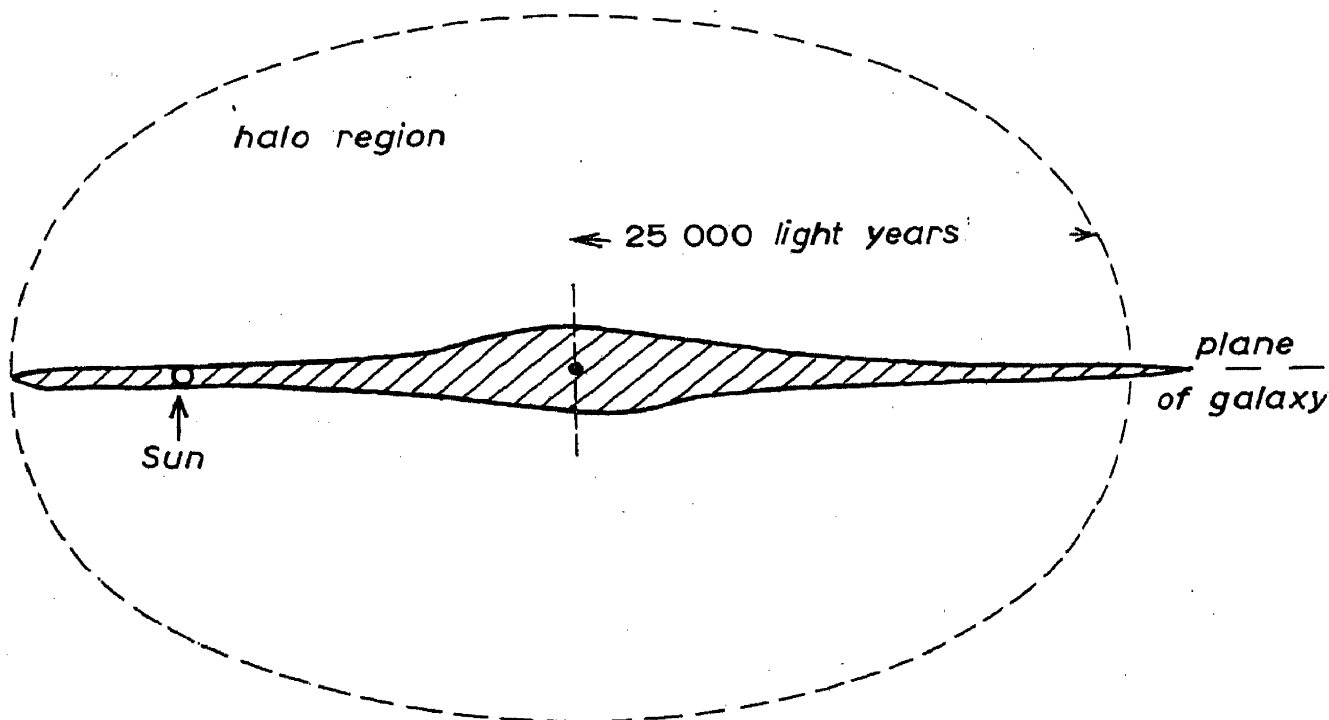


Fig. 2. - Schematic cross-section through the Galaxy.

but the rotational velocity at a certain distance from the centre depends on the gravitational attraction of the whole system. At great distance from the centre Kepler's laws are approximately valid but the rotation curve further in is more complicated because the gravitational field cannot be approximated by that of a point mass at the centre of the system. The best values of the rotation velocities at points within the Sun's orbit come from observations of the 21 cm hydrogen line.

Some classes of stars and particular objects like the globular clusters are also observed at great distances from the plane of the disk. They are found in a nearly spherical volume surrounding the disk (Fig. 2), commonly called the galactic « halo ». Only old stars are found there, namely, the stars formed before the galactic gas had contracted into a disk. The forces of fluid dynamics and magneto-fluid dynamics which gave the gas in the disk its present structure, have no hold on the stars after their formation. The halo stars thus are « fossils » of an early dynamic stage in galactic evolution.

Photographs of other galaxies show a spiral structure in the disk and this is the case also in our galaxy. However, it is extremely difficult to study this structure optically because of the absorbing dust clouds in the plane of the

disk. Our knowledge about the spiral arms in the galaxy away from the immediate neighbourhood of the sun is entirely based on observations of the 21 cm hydrogen line emission.

In trying to describe the local dynamics of the gas in the disk, say, somewhere in a spiral arm, we immediately run into a peculiar situation. By careful estimates of the gains and losses of energy (heating and cooling mechanisms) SPITZER and SAVEDOFF found the temperature of the gas in the galactic disk to be 10^2 °K in the regions where the hydrogen is neutral, and 10^4 °K where it is ionized. The low temperature of neutral hydrogen has since been confirmed by radio measurements. But the dynamical consequences of this factor 100 between the temperatures in the two regions were not at once realized. It is obviously immaterial whether this is precisely 100, or 50, or 200.

A star that forms in an initially neutral gas cloud ionizes the gas out to a certain distance producing the «Strömgren sphere» discussed earlier. If the densities inside and outside are about the same initially, the kinetic temperature and hence the pressures will differ by a factor 100. This is an unstable situation and an explosive expansion of the gas surrounding the young star will follow. Also individual clouds in the vicinity may be «rocketed» away from the star. On the side where the cloud faces the star the gas is being ionized and heated to 10^4 °K. This hot part evaporates very fast and the «recoil» sends the remaining part of the cloud away from the star. This is a genuine cosmic rocket driven by the nuclear power which is produced, away from the rocket, in the body of the hot star.

The general effect of these expansion and rocketing processes is to stir up the interstellar gas and give it the irregular aspect which is so striking on all photographs. Besides young, hot stars, supernova explosions will play a similar role.

The question of the physical conditions in the galactic halo is extremely difficult. The strongest arguments for the existence of a halo come from observations of the continuous radio frequency radiation from high galactic latitudes. There are also a few indirect arguments derived from optical observations.

Statements about the temperature and density in the galactic halo have to be mainly guess work: a possible combination is $T=10^6$ °K and $N_e=10^{-4}$ cm⁻³, which would make it an approximately self-supporting «atmosphere». Such a gas is impossible to observe with present techniques. An intimate relation between the dynamics of the halo and the disk gas is indicated by the expansion found in the inner parts of the disk. Careful 21 cm measurements made in Dwingeloo, Holland, have revealed that the gas in the disk within 3 kps from the centre streams away from the centre, while also partaking in the rotation. The loss of one solar mass per year in the central region can only be compensated by downpour from the halo.

The forces acting in these processes are very poorly understood but it is generally believed that magnetic fields play an important role.

The theoretical case for such fields is rather clear. Magneto-hydrodynamics and specially the inductive coupling of gas and magnetic fields is of great importance anywhere in astrophysics. The «lifetime» of a magnetic field in a conducting medium is proportional to the square of the linear dimensions times the conductivity. The conductivity of the interstellar gas is comparable with that of ordinary metals but the dimensions are so large that it is difficult to see how the field can leak away at all in the time of a thousand million years. In contrast, these phenomena are difficult to study in the laboratory because it is hard to maintain the fields longer than a microsecond. The field is probably to some extent lined up along the spiral arms because the observed polarization of the starlight is greater and more consistent when looking across the local spiral arm than when looking along it. The recent radio polarization measurements in Holland seem to show that the radio polarization vector is perpendicular to the optical. This is to be expected as synchrotron radiation is polarized perpendicular to the field, while the interstellar optical polarization, which is a completely different process, results in polarization parallel to the field. However, these first sample measurements are in urgent need of further confirmation.

The dynamics of the spiral arms, in particular their maintenance against the distortion produced by differential rotation, is still imperfectly understood. Three factors may be at work: the peculiar gravitational potential, which is responsible for differential rotation; the magnetic fields; and the explosions described above which may blow gas from one arm to another and may produce a turnover of the complete gas content of an arm in 10^9 year.

As regards the magnetic field in the halo, I can say very little; practically all our ideas about this come from various cosmic ray theories. I do not consider it appropriate to enter here into a discussion of topics which will be treated more fully by other speakers and I hope that this very brief summary of some facts and methods will be useful to you in your further discussions.

The Polarization of Starlight in the Interstellar Medium (*).

J. M. GREENBERG (**)

Rensselaer Polytechnic Institute - Troy, N.Y.

Introduction.

Since the discovery of the polarization of starlight (HALL [1]; HILTNER [2]) by the interstellar medium, the most commonly accepted theoretical explanations have involved the scattering by grains which are oriented by galactic magnetic fields. In the sense that no other orientation mechanism has yet been shown to be adequate it may be claimed that the existence of polarization is evidence for these magnetic fields and, as a matter of fact, is perhaps the most direct evidence for the large scale galactic magnetic fields. Because the theories of the acceleration and trapping in the galaxy of high energy particles are based on interactions of these charged particles with magnetic fields it is appropriate that the subject of interstellar polarization should be discussed here. There has been a time when people working in cosmic rays or in interstellar polarization used each other's theories to prove the existence of the required magnetic fields. This was, of course, not a very satisfactory situation and we will attempt here to show that it may be possible in the near future to base independent and quantitative inferences on the distribution of galactic magnetic fields on the phenomenon of interstellar polarization.

Since polarization occurs only when there is extinction of the starlight it is generally assumed that both phenomena are produced by the same basic process; namely scattering and absorption of radiation by small grains. The present state of astronomical observations and theories does not yet definitely establish the nature of the interstellar grains. In order to give a complete story one should try to consider all possibilities equally. However this can not be

(*) Work supported by the National Science Foundation and the National Aeronautics and Space Agency.

(**) At present on leave from Rensselaer Polytechnic Institute as an O.E.E.C. Senior Visiting Fellow at Leiden Observatory.

done because some of the theories have been much further developed than others and therefore the greater emphasis is inevitably placed on these.

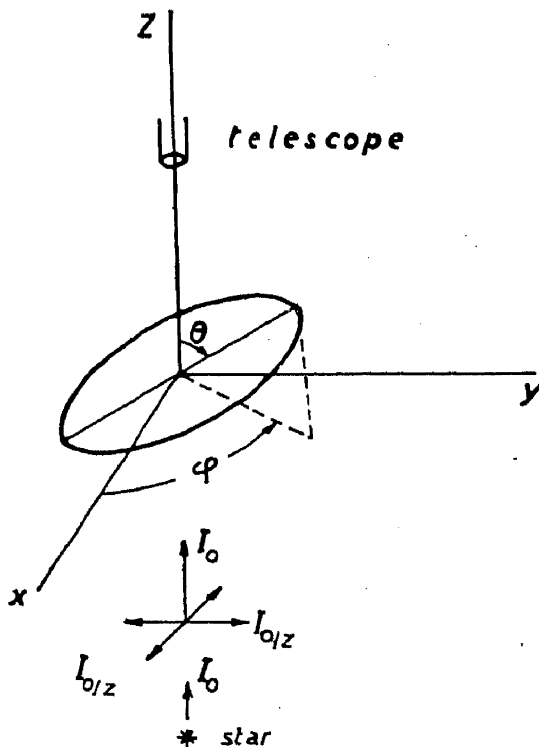
In the first section there are summarized the characteristic data on extinction and polarization and a general, though somewhat simplified, formulation of the means by which these are produced by scattering is presented. The three fundamental theoretical problems which are considered in the following two sections are: 1) the physical and chemical nature of the grains, 2) orientation of the grains, and 3) scattering of light by the grains. In the last section the predictions based on the various types of grains are compared.

1. - Extinction and polarization.

The radiation from a star is reduced by a combination of scattering and absorption by all of the grains between the stars and the observer. The total (extinction) cross-section by a single grain is denoted by σ and is defined by

$$\sigma = \frac{\text{Radiant energy scattered and absorbed per unit time}}{\text{Incident radiant energy per unit area per unit time}}$$

In general σ depends on the wave length of the radiation and the orientation of the particle. For simplicity, let us assume axially symmetric particles, all having the same degree of nonisotropy of shape (same elongation). The extension to the general case is easily made. Let the direction of the line of sight be denoted by the z -axis (see Fig. 1) and divide the incident radiation into two incoherent polarized beams, each of intensity $I_0/2$, oriented along the x and y axes. Then, at the observer, the star light is attenuated according to



$$(1) \quad \begin{cases} I_x = \frac{I_0}{2} \exp[-\tau_x], \\ I_y = \frac{I_0}{2} \exp[-\tau_y], \end{cases}$$

Fig. 1. - Geometric configuration of an interstellar grain in the line of sight between star and observer.

where the optical depths τ_x and τ_y are given by summing the cross-sections of all the

particles along the line of sight

$$(2) \quad \begin{cases} \tau_x(\lambda) = \int_0^L \int_{\text{angles}} \int_0^\infty n(a) \sigma_{E_x}(a, \theta, \varphi, \lambda) f(a, \theta, \varphi) d\Omega dZ, \\ \tau_y(\lambda) = \int_0^L \int_{\text{angles}} \int_0^\infty n(a) \sigma_{E_y}(a, \theta, \varphi, \lambda) f(a, \theta, \varphi) d\Omega dZ, \end{cases}$$

where L = path length, $n(a)$ is the function giving the distribution of the particles in size (represented by the parameter a), $f(a, \theta, \varphi)$ is the angular distribution function (which will generally be size-dependent), σ_{E_x} and σ_{E_y} are the cross-sections for light polarized along the x and y directions. The total received intensity is given by

$$(3) \quad I = I_x + I_y = \frac{I_0}{2} [(\exp[-\tau_x] + \exp[-\tau_y])],$$

and the attenuation, expressed in magnitudes, is

$$(4) \quad \Delta m = -2.5 \log \frac{I}{I_0} \approx \frac{1.086}{2} (\tau_x + \tau_y),$$

assuming $(\tau_x - \tau_y)/(\tau_x + \tau_y) \ll 1$, which is generally a good approximation. The polarization is given by

$$(5) \quad p = \frac{I_x - I_y}{I_x + I_y} \approx \frac{\tau_x - \tau_y}{2} \quad (p \ll 1).$$

It is often convenient to express the polarization as a magnitude difference

$$(6) \quad \Delta m_p = -2.5 \log \frac{I_x}{I_y} = 1.086 (\tau_y - \tau_x) \approx 2 \times 1.086 p.$$

It is seen that both the extinction and polarization in magnitudes are linearly related to the extinction cross-sections of the grains.

The basic character of the experimental results on the polarization and the wave length-dependence of the extinction are schematically illustrated in Fig. 2. The difference between the upper and lower curves gives the polarization. The roughly linear region for $3000 \text{ \AA} \leq \lambda \leq 10000 \text{ \AA}$ is often referred to as the λ^{-1} law.

The maximum observed value of the ratio of polarization to extinction implies $\tau_y/\tau_x \leq 1.06$.

The largest polarizations are generally observed for stars in the galactic plane. HALL [3] has compiled a vast amount of data on polarization and has constructed maps showing the degree and orientation of the polarization. From

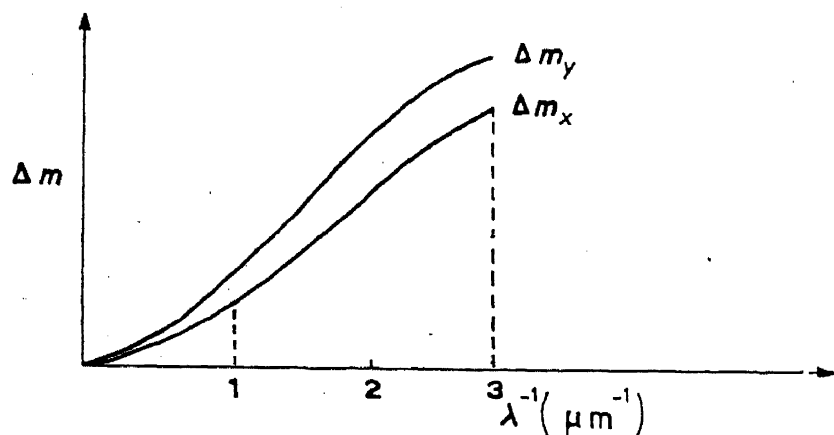


Fig. 2. — Schematic observational data on interstellar extinction for two orthogonal directions of polarization. In the absence of interstellar polarization the upper and lower curves would be identical.

a broad point of view it is seen that in some extended regions of galactic longitude the polarization is parallel to the plane of the galaxy and that the dispersion of the polarization angles is quite small ($\sim 5^\circ$), whereas in other regions the planes of polarization are random. The regions of uniformity appear to correspond to stars seen across a spiral arm and the regions of high dispersion appear to correspond to stars viewed along a spiral arm. On a much

smaller scale, there is found to be a definite tendency for the polarization (electric vector) to be parallel to nebular streamers, striated clouds and filamentary structures. The wave length-dependence of the polarization is not well known, although there generally appears to be a maximum somewhere between the infrared and ultraviolet (GEHRELS [4]; BEHR [5]). The most direct evidence for the orientation of interstellar grains by magnetic fields will be obtained if the preliminary results on the correlation between radio and optical polarization (SEEGER, WESTERHOUT, TINBERGEN and BROUW [6]) are confirmed.

2. — Physical properties of proposed grains and their orientation.

The types of grains which have been proposed and a summary of their chemical and physical characterization are presented in Table I.

All of the theories using these grains to account for extinction or polarization require a total amount of any element which is at least consistent with the normally stated cosmic abundances of the elements. The size given is the characteristic average linear dimension required to account for the wavelength dependence of the interstellar extinction. The scattering properties of all but the trapped free radicals are given by the classical application of the scattering of electromagnetic radiation, and will be discussed in some detail in the next section. The trapped free radicals scatter by a quantum mechan-

ical process of resonance absorption and subsequent reemission and only qualitative information is available for this process.

TABLE I.

Type of particle	Important chemical constituents	Magnetic character	« Size » a	Scattering
Dielectric (dirty ice)	H, O, C (metallic trace)	Para-magnetic	$0.5 \mu\text{m}$	Real refractive index (Fig. 3)
Metallic	Fe, Mg	Ferro-magnetic	$0.05 \mu\text{m}$	Complex refractive index (Fig. 4)
Graphite flakes	C	Dia-magnetic	$0.05 \mu\text{m} \div 0.5 \mu\text{m}$	Dielectric for $E_{\parallel}$ c -axis; Metallic for $E_{\perp}$ c -axis (Fig. 5)
Trapped free radicals	H, O, C	Para-magnetic	$10 \text{ \AA}$ $\left(\frac{a}{\lambda} \approx \frac{1}{400}\right)$	Resonance absorption (Quantum mechanical)

The orientation mechanism which will be presented here in very abbreviated form is that due to DAVIS and GREENSTEIN [7].

When a non spherical particle spins in a magnetic field, the field as seen by the particle appears to vary sinusoidally. Because the magnetic susceptibility has an imaginary part which is frequency-dependent, there is a phase difference between the magnetization and the external field. This gives rise to a small nonconservative torque which tends to orient an axially symmetric particle so that it rotates about its short axis. Collisions of the gas atoms with the grains tends to disorient them. The relaxation time for the orientation is given by (in c.g.s. units)

$$(7) \quad \tau_B \approx \frac{\omega I \gamma}{\chi'' V B^2},$$

where ω is the angular velocity of the particle, I is its moment of inertia, γ is a measure of non-sphericity, V is the volume, B is the magnetic field and χ'' is the imaginary part of the susceptibility. The relaxation time for disorientation is (in c.g.s. units)

$$(8) \quad \tau_{\text{coll}} \approx \frac{5 \cdot 10^{19} \rho_g a}{n_H T^{\frac{1}{2}}},$$

where ρ_g is the density of the grains, a is a characteristic linear dimension, n_H is the number density of hydrogen and T is the gas temperature.

In order to achieve a high degree of orientation τ_{coll} should be much greater than τ_B . Assuming a reasonable value of χ'' for the dielectric grains and using accepted values for the astrophysical quantities involved in the formula ($n_B T^{\frac{1}{2}} = 100$) gives for the required value of the magnetic field

$$B > 2 \Gamma \quad (1 \Gamma = 10^{-5} \text{ gauss}).$$

This is somewhat high but perhaps not unreasonably so. The need for a more precise calculation of this method of orientation than that of DAVIS and GREENSTEIN would be very helpful because even in their paper a number of serious approximations are made and it is not readily seen whether the magnetic field requirement is underestimated or overestimated. The value of the magnetic field required for orientation of the other types of grains may be similarly computed and one can show that for the metallic and trapped free radical particles a lower value of the field is needed. However, several fundamental questions involving the internal temperatures of such grains lead to difficulties and uncertainties regarding the interpretation of such values.

3. - Scattering.

Figure 3 shows some smoothed experimental results using the microwave analog method (GREENBERG, PEDERSEN and PEDERSEN [8]) to obtain the

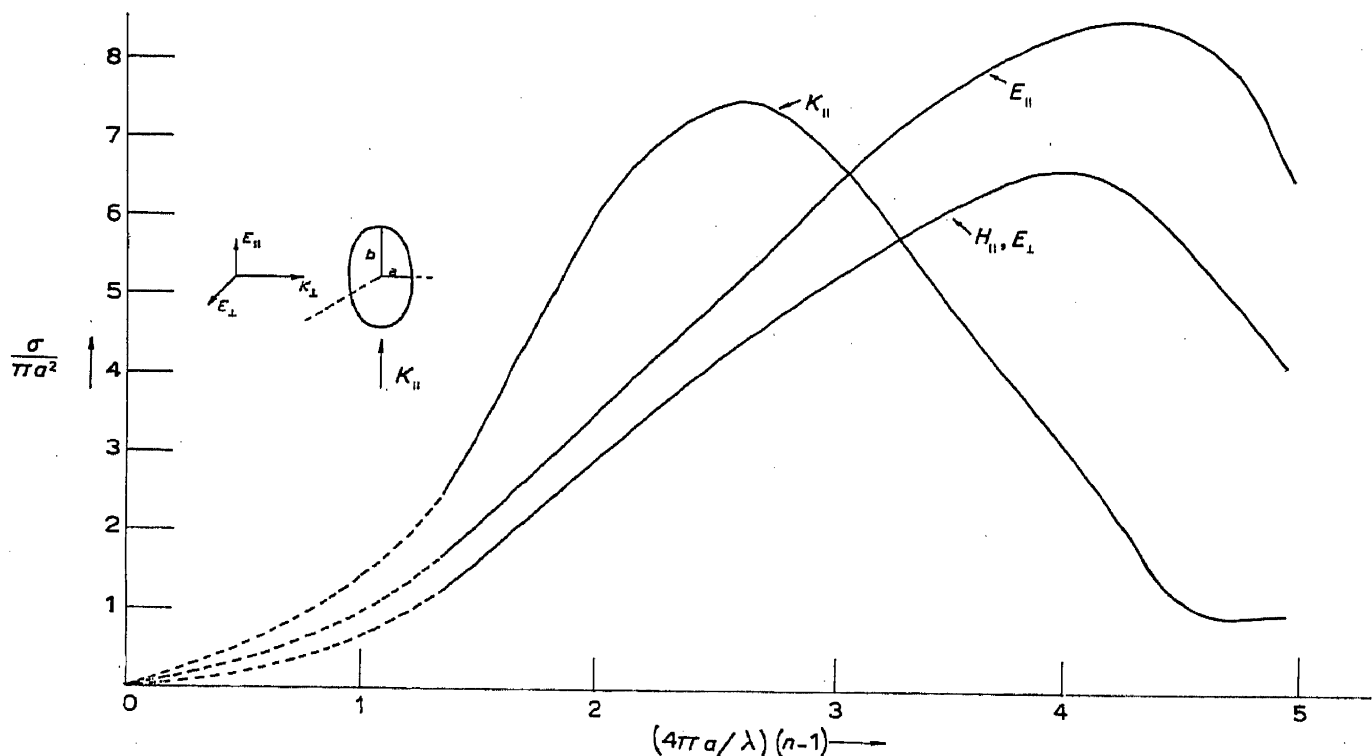


Fig. 3. - Smoothed experimental results for the total cross-sections of spheroidal particles: $n=1.26$, $b/a=2$. The dotted curves are extrapolations of the experimental curves. The labels $K_{\parallel}$, $E_{\parallel}$, $E_{\perp}$, $H_{\parallel}$ refer to the angles of the propagation, electric, and magnetic vectors of the radiation with respect to the symmetry axis of the particle.

scattering by spheroidal particles of elongation $b/a = 2$, refractive index $n = 1.26$, and for three orthogonal orientations of the particles relative to the incident radiation. It is clear that a cloud of such particles spinning with their short axes parallel to the spiral arm axis will produce the desired direction of polarization. It is generally inferred from this that magnetic fields tend to line up with the spirals arms. Furthermore it can be shown qualitatively that the scattering by a distribution of sizes of such spheroidal particles which reproduces the extinction law will give a value of $\tau_y/\tau_x \approx 1.1$ for perfect orientation. This is perhaps a sufficient margin to allow for imperfect orientation and still obtain the value needed ($\tau_x/\tau_y \approx 1.06$) to account for the maximum amount of polarization observed. The characteristic particle size

(the single-particle size which will roughly produce the extinction law) can be obtained by letting $\lambda \approx 5000 \text{ \AA}$ and by choosing that value of a which corresponds to the middle of the linear region of the curves. The scattering by metallic particles ($n \approx 1.3(1 - i)$) is shown schematically in Fig. 4. The char-

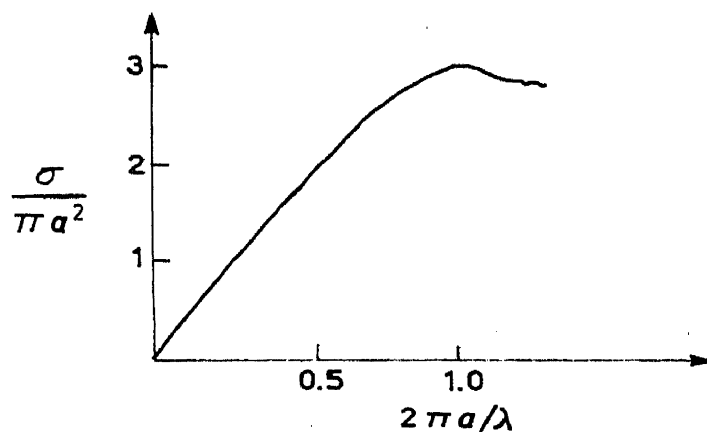


Fig. 4. — Characteristic variation of the cross-section of a metallic particle ($n \approx 1.3(1 - i)$) with size.

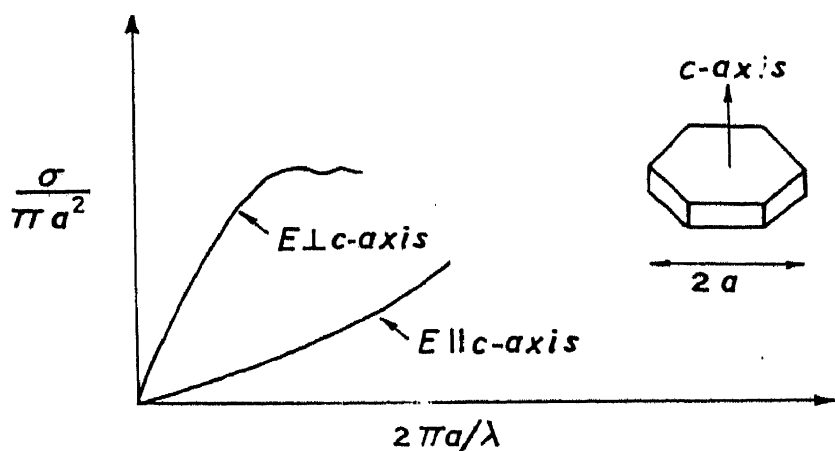


Fig. 5. — Hypothetical scattering characteristics of a graphite flake for two orthogonal orientations of the c -axis with respect to the electric vector of the incident radiation.

acteristic size of the metallic particles can be seen to be about $1/10$ that for dielectric particles. Because there is not yet sufficient quantitative information on the scattering by nonspherical metal particles in this size range it is possible only to state that they should at least be adequate polarizers. The crystal structure of graphite particles implies that for the electric vector of the radiation parallel to the c -axis (axis perpendicular the flakes) they act like dielectrics and

for the electric vector perpendicular to the c -axis they act like conductors. On this basis there are given the schematic curves of Fig. 5 and it is easy to see why such grains should be excellent polarizers even when imperfectly oriented. The trapped free radical particles have been shown by

PLATT [9] to scatter by a quantum-mechanical resonance at a wave length $\lambda \approx 400 \alpha$ and also to be effective scatterers at shorter wave lengths. A quantitative treatment of the process by which these particles can polarize the star light is not possible. However, qualitative calculations (PLATT [9], GREENBERG [10]) seem to indicate that they could be sufficiently effective polarizers.

4. - Concluding remarks.

Existing observational data (some of which has not been included here) appear to imply that the interstellar grains are most likely to be either of the dielectric type or trapped free radical type.

The others were included in the discussion primarily to show that with the assumption of the Davis-Greenstein mechanism for orientation the directions of the magnetic fields are similarly defined in all cases but that there still remains the open question of their magnitudes. Finally, there now appears to be well defined ways of making a unique selection of the particle type based on further astrophysical observations of the wave length dependance of the excitation and polarization (GREENBERG and MELTZER [11]). When this has been accomplished the problem of calculating the strength of the magnetic fields from the degree of polarization will be much closer to solution.

REFERENCES

- [1] J. S. HALL: *Science*, **109**, 166 (1949).
- [2] W. A. HILTNER: *Science*, **109**, 165 (1949).
- [3] J. S. HALL: *Publ. U. S. Naval Obs.*, vol. **17**, part 6 (1958).
- [4] TH. GEHRELS: *A. J.*, **65**, 470 (1960).
- [5] A. BEHR: *Zeits. f. Astrophysik*, **47**, 54 (1959).
- [6] C. L. SEEGER, G. WESTERHOUT, J. TINBERGEN and W. BROUW: *B. A. N.*, **16**, in press (1961).
- [7] L. DAVIS, Jr. and J. L. GREENSTEIN: *Ap. Journ.*, **114**, 206 (1951).
- [8] J. M. GREENBERG, N. E. PETERSEN and J. C. PEDERSEN: *J. A. P.*, **32**, 233 (1961).
- [9] J. R. PLATT: *Ap. Journ.*, **123**, 486 (1956).
- [10] J. M. GREENBERG: *Ap. Journ.*, **132**, 672 (1960).
- [11] J. M. GREENBERG and A. S. MELTZER: *Ap. Journ.*, **132**, 667 (1960).

For further discussion on interstellar polarization and a more complete bibliography see *Lowell Observatory Bulletin*, No. 105, vol. 4 (1960), p. 264-321.

Optical Astronomy above the Earth's Atmosphere.

J. M. GREENBERG (*)

Rensselaer Polytechnic Institute - Troy, N. Y.

Introduction.

In order to picture the possible impact of observations involving large new bands in the electromagnetic spectrum it is only necessary to consider the obvious example of the effects of the advent of radio astronomy on our understanding of our galaxy and the universe. A further analogy may be made with the value of the early studies of cosmic radiation in providing an energy spectrum of nuclear interactions which led to important advances in nuclear physics. Unfortunately the analogy ends here in the sense that whereas laboratory work through high energy accelerators can successfully imitate the phenomenon of cosmic rays, in astronomy we are largely limited to being passive observers and can generally only theoretically extrapolate to, rather than duplicate in the laboratory, the conditions and phenomena which we hope to explain.

No attempt will be made here to present a complete survey of the important problems which present themselves with the advent of the various proposed observations throughout the electromagnetic spectrum. Some of these have already been discussed here and, of course, there already exists a large literature of summaries and detailed studies of specific problems.

We will however, try to cover some topics which are illustrative of the main features of the present studies and we will try to cover in somewhat greater detail those fields which appear most closely to relate to the general title of this course. Furthermore we will almost exclusively consider only the ultra-violet spectrum below the atmospheric cut-off by ozone (2900 Å).

As a means of getting oriented in the subject it is useful to see the limitations which the earth's atmosphere places on astronomical observations.

(*) At present on leave from Rensselaer Polytechnic Institute as an O.E.E.C. Senior Visiting Fellow at Leiden Observatory.

These limitations are schematically shown in Fig. 1 [1]. The figure not only shows the atmospheric absorption but also shows such things as the maximum of the blackbody radiation as a function of temperature. It can be seen that for $T \sim 10\,000^\circ\text{K}$ the maximum falls at about the edge of the short wavelength limit of the optical window.

Since temperatures considerably higher than this occur with high frequency it becomes immediately apparent that any studies which involve considerations of such temperature are severely hampered. It is worthwhile to consider

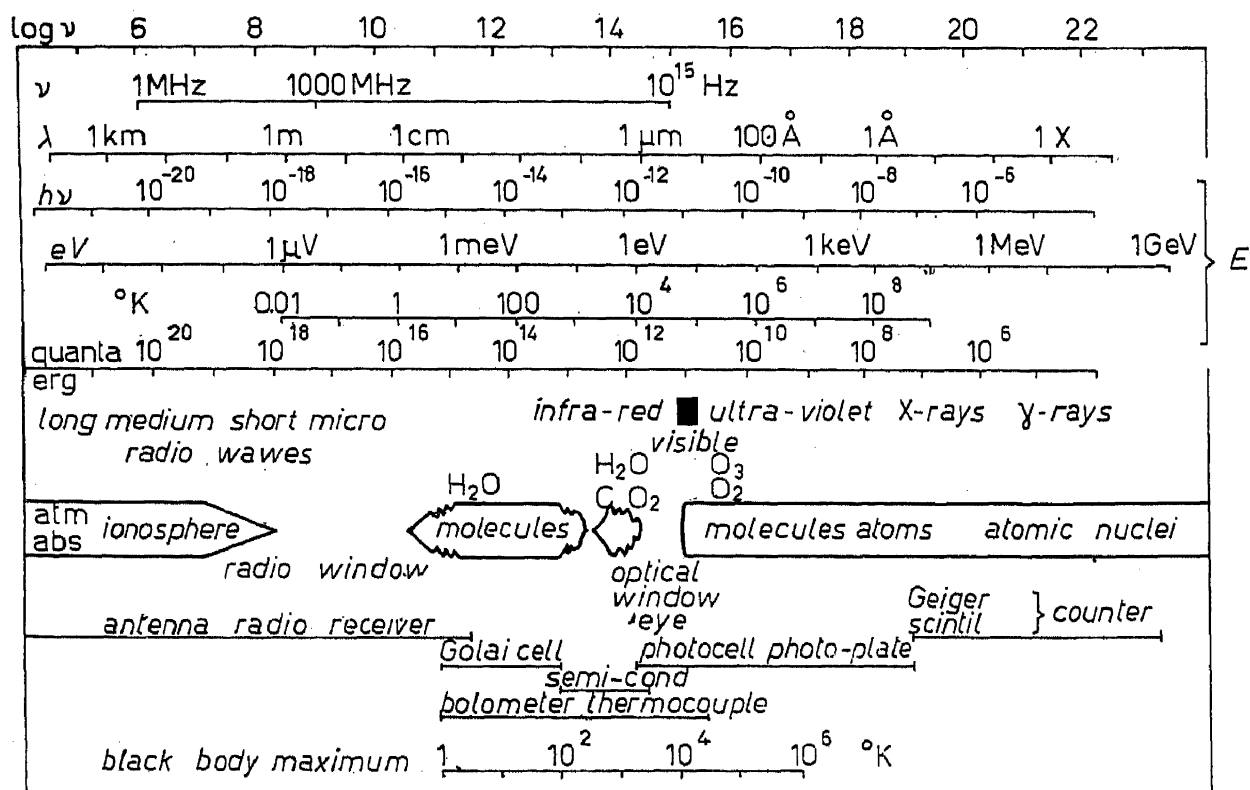


Fig. 1. — Schematic representation of the transparency of the terrestrial atmosphere as a function of wave length [1].

some of the already existing results of far ultraviolet work which has been done using rockets. Although this technique is subject to limitations it will play an extremely important role in preliminary surveys which will be used to prepare the way for the design of the satellite observatories. HINTEREGGER [2] has presented some preliminary results of EUV (extreme ultraviolet) fluxes from the sun (see Fig. 2). We are not so interested for the moment in the quantitative aspects of these results but rather to show that even though the photosphere temperature of the sun is only $6\,000^\circ\text{K}$ nevertheless we observe spectral features corresponding to extremely high temperatures, ($T \sim 10^6^\circ\text{K}$), and that, as is well known the resonance lines of the hydrogen and helium lie in a region which has heretofore not been observationally avail-

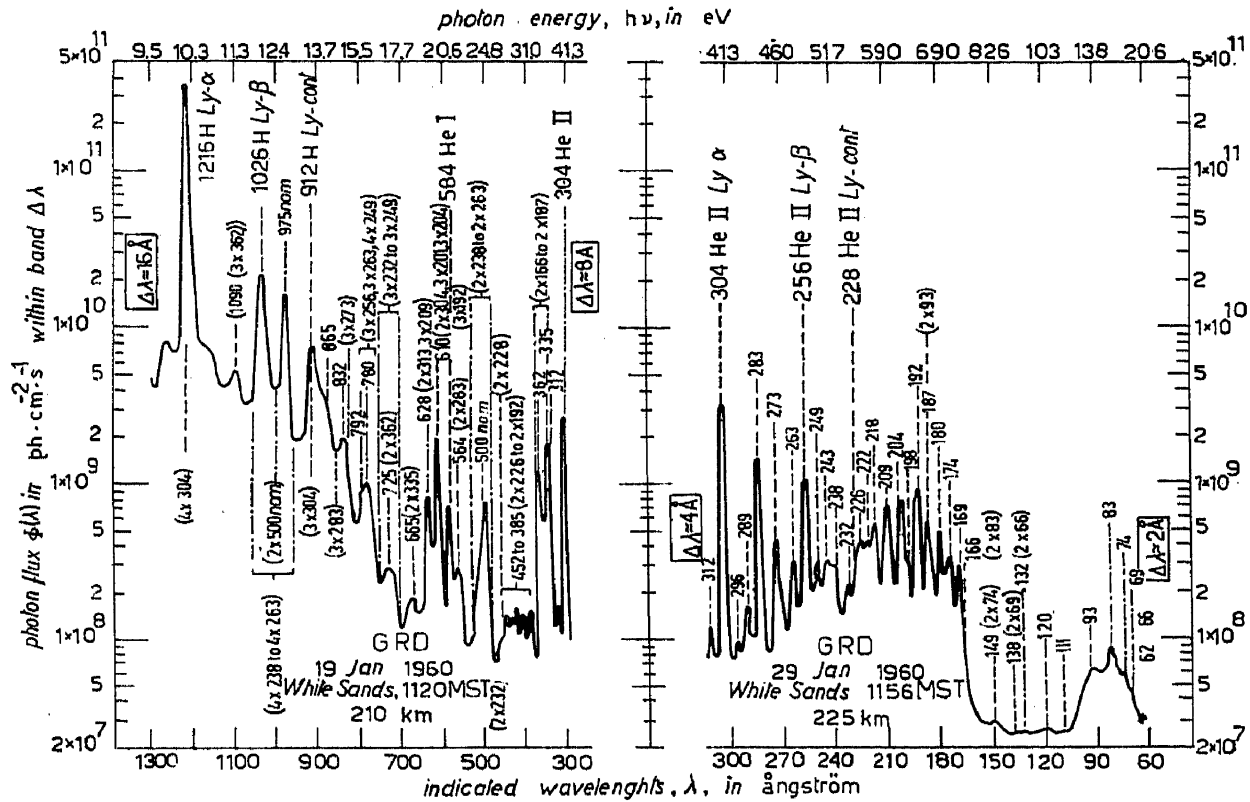


Fig. 2. - Solar extreme ultraviolet fluxes in the upper atmosphere as determined from rocket flights.

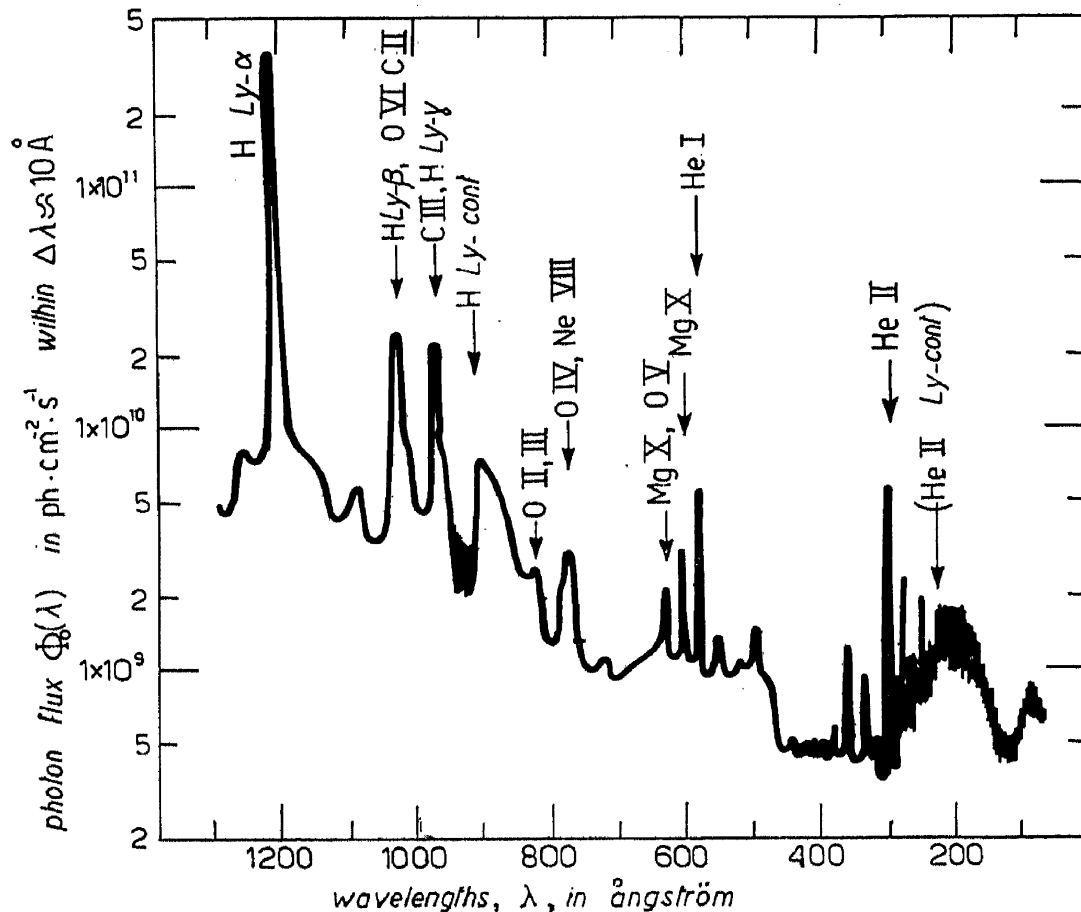


Fig. 3. - Estimated solar EUV photon fluxes on top of the earth's atmosphere. The ordinates represent approximate photon-flux densities for an assumed constant wavelength bandwidth of 10 Å .

able. Figure 3 shows a graphical representation [2] of the photon fluxes between 1300 and 60 Å that one may expect to measure with an instrument of spectral bandwidth 10 Å placed above the earth's atmosphere. As is well known the excitation to highly ionized states of the heavier elements occurs only in the very high temperature of the chromosphere and corona.

One of the fascinating problems which will undoubtedly receive attention is the question of whether the same temperature inversion seen in the sun's atmosphere occurs in other stars. Also, how different will the coronas of stars of different spectral type be. This may be intimately related to the problem of the injection of high energy particles into interstellar space. To show where some of the observed and expected resonances lines from the sun are, we present Table I [3]. The need of such calculations for identification purposes is stimulating renewed and expanded efforts in the field of theoretical spectroscopy. There is a singularly important point with regard to stellar spectroscopy which should be kept in mind, and this regards the fact that just as our own atmosphere presents an obstacle to radiation reaching us, so also does the matter in interstellar space. We postpone a general discussion of the positive aspects of observation of this interstellar matter but mention here the negative role that hydrogen and helium are expected to play in obscuring the radiation of almost all distant stars. ALLER [4] has summarized his findings in the line and continuous absorption by hydrogen and finds for example, that for 10^{21} atoms/cm², the optical depth τ would be 6000 at the Lyman limit, about 6 at 100 Å and about 0.005 at 10 Å. Assuming a hydrogen to helium ratio of 7, the optical depth due to He at $\lambda = 504$ is 1000. At higher frequencies the helium absorption is more and more effective relative to the hydrogen absorption and the net effect is that the strong Lyman continuum will effectively extinguish all the radiation between $\lambda = 912$ and the soft X-ray region except perhaps in the spectra of some of the nearest stars and those at high galactic latitudes.

1. - Continuous spectra of the stars.

The distribution of stars according to their temperature and luminosity is illustrated in Fig. 4 [5]. The temperatures of very hot blue stars of types O—B—A have been estimated from their line and continuous spectra to range from 12000 to 50000 °K. A knowledge of the radiation of such stars in the ultraviolet is important because these hot bright stars are the chief sources of radiation at least in the outer parts of our galaxy. The temperatures of interstellar gas and dust are controlled largely by the energy density of this dilute radiation. Furthermore any estimate on the life-times and evolutionary processes of such stars is based on their rates of energy produc-

tion. Figure 5 [1]) shows the energy distribution calculated for a main sequence G0 stars and a B0 star compared with observations. While the observations appear to fit quite well the theoretical prediction in the usual optical range, it is quite evident, as shown in Fig. 6 [1], that we could be in

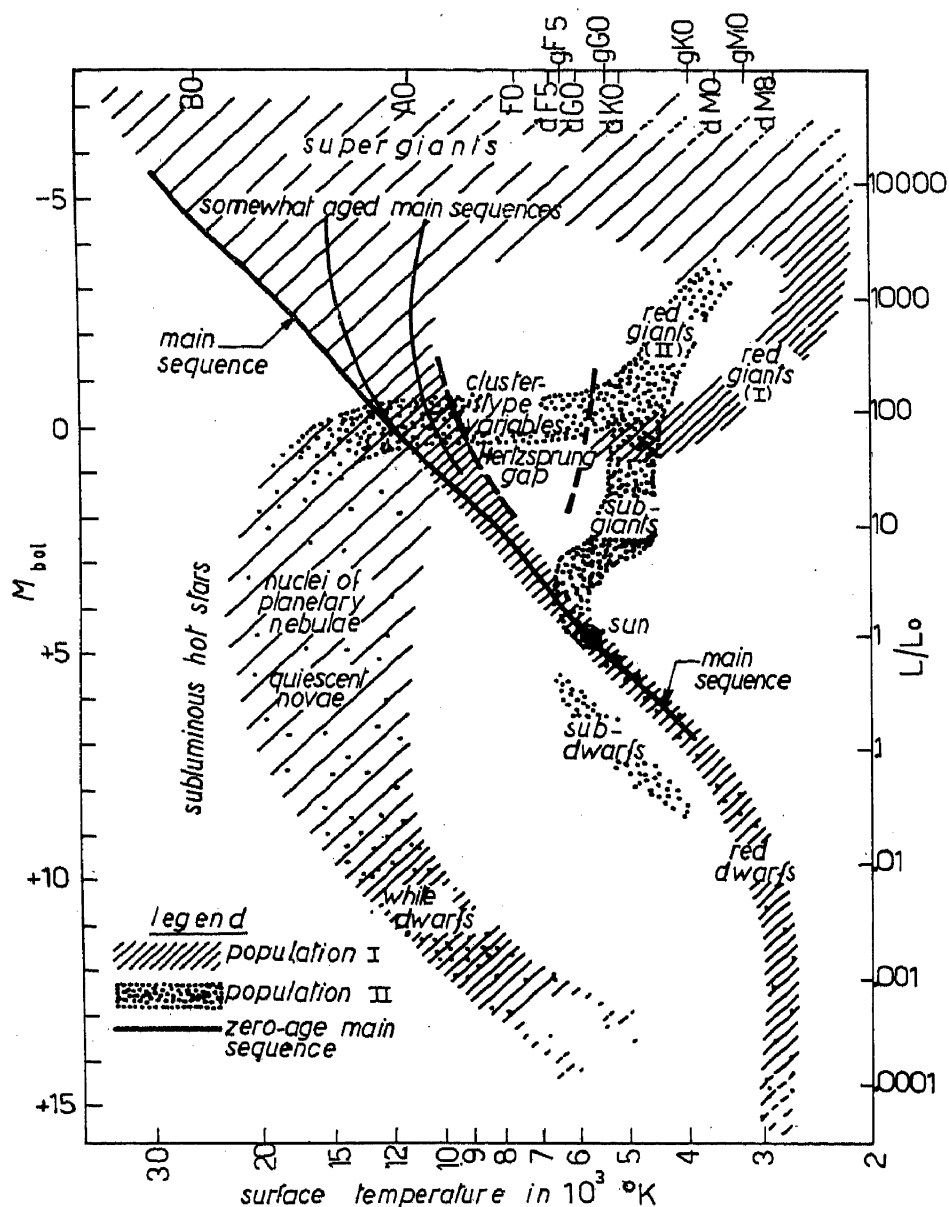


Fig. 4. —Hertzsprung-Russell or temperature-luminosity diagram, showing the regions occupied by stars of various kinds of populations.

for many surprises in the region of extrapolation into the far ultraviolet. The early work by the Naval Research Laboratory has, as a matter of fact, already shown this to be true. BOGGESS reports that the stellar flux in the band $\lambda = 1225$ to $1350 \text{ \AA}$ in a B0.5 V star (ϵ Persei) was about a factor of thirty less than predicted [6].

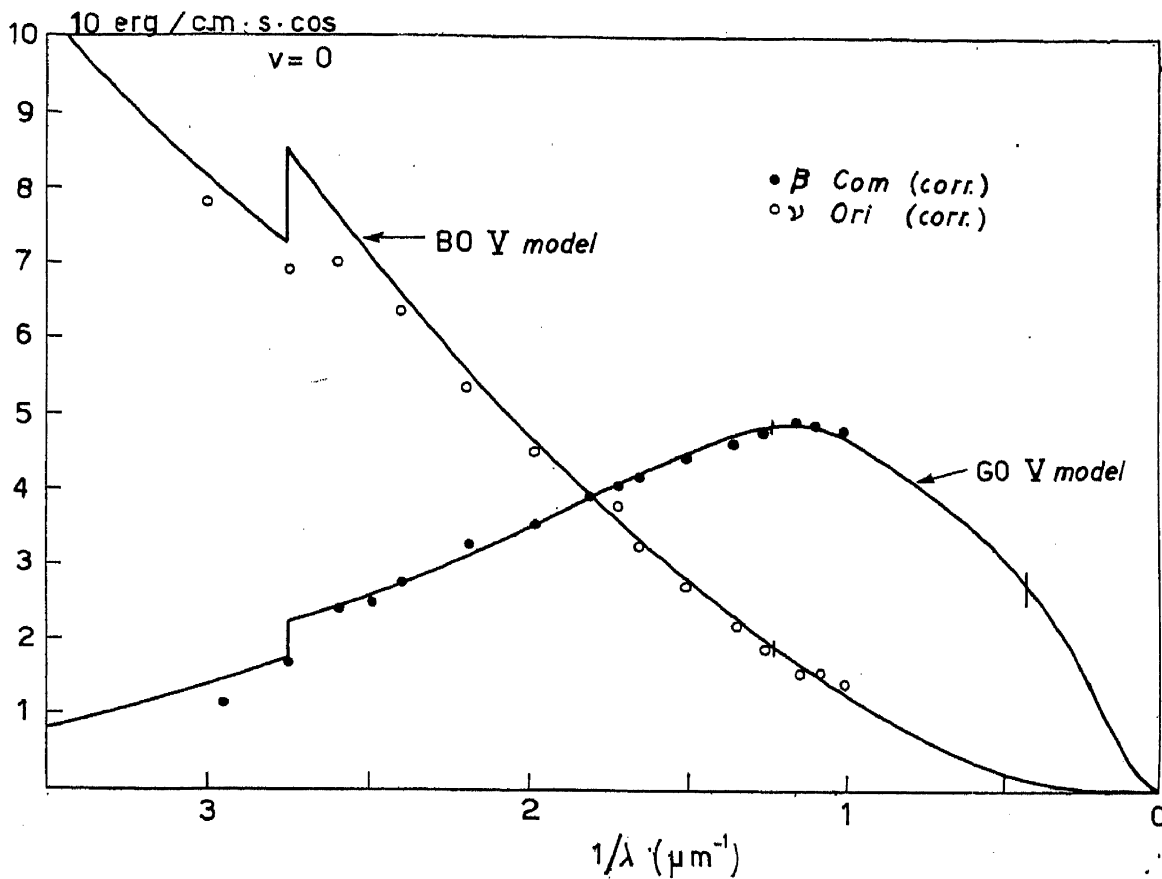


Fig. 5. - Observed spectral distribution of the energy radial in the wavelength interval $3000 \div 10\,000 \text{ \AA}$ ($0.3 \text{ }\mu\text{m}$ to $1.0 \text{ }\mu\text{m}$) by stars of type G0 V ($\sim 6\,000 \text{ }^\circ\text{K}$) and B0 V ($\sim 25\,000 \text{ }^\circ\text{K}$).

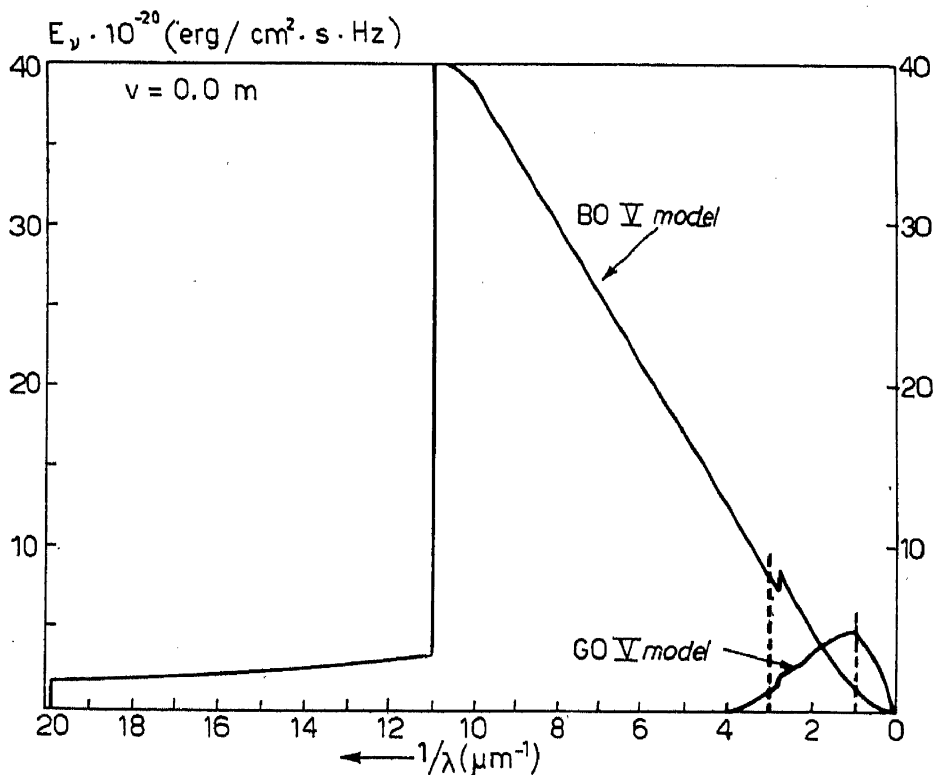


Fig. 6. - Spectral distribution of the radiation from stars of type G0 V and B0 V, extrapolated into the ultraviolet from the observed data of Fig. 5.

2. - Stellar spectroscopy.

As has already been shown, the resonance lines of the most common elements occur in the ultraviolet (see Table I). Abundance determinations from ground-state transitions are the most reliable. These abundance determinations are one of the distinguishing features between Type I and Type II stars. It is expected that the Type I stars which occur in the spiral arms are being formed out of the gas and dust and analysis of the absorption line spectra of stellar atmospheres by the curve-of-growth technique has indicated that these stars have a generally higher ratio of heavy elements to hydrogen than do the type II stars. The curve of growth is theoretically determined as a plot of the equivalent width of an absorption line as a function of the product of the number of atoms times the f (oscillator strength) value, taking in account the Doppler effect. If we know the shape of a spectral line and consider the corresponding absorption through the stellar atmosphere of the continuum spectrum of the star, the abundance of a given element can be derived by the kink appearing in the continuum spectrum up to a limit which is reached when the blacking out is complete.

3. - The interstellar medium.

Observations of atomic absorption lines should give detailed information on the physical conditions of the interstellar gas—its temperature, level of ionization, electron density and chemical composition. The only interstellar absorption lines of any strength in the presently accessible spectral region are those of NaI and CaII (hydrogen is about 10^5 to 10^6 times as abundant as sodium). The newly opened spectral range will make available the lines of more abundant atoms such as C, N, O, Mg, Fe in different states of ionization. It will be possible to make great improvements in the observation of the velocities of neighbouring gas clouds and thus to have a much better foundation for theories of interstellar gas dynamics. Insofar as the density of hydrogen is concerned, we have here a problem created by its great abundance. The absorption by neutral hydrogen will very often be so high that the Lyman α line will be completely saturated over several angstroms width. However simply by going up the scale of excitation it may still be possible to consider sufficiently weak lines to obtain direct quantitative estimates of the hydrogen density. In neutral hydrogen regions the 21 cm work has given the only direct information on hydrogen density. In regions where hydrogen is predominantly ionized the hydrogen density has had to be inferred from rather rough estimates of absorption by relatively uncommon substances (Ca, Ca⁺, Na, K⁺, Ti⁺) taken together with an assumed hydrogen/metal abundance ratio.

Now, in principle, measurements of the strengths of the higher members of the Lyman series of hydrogen together with measurements of the strength of the Fe II lines ($(2\,000 \div 2\,700) \text{ \AA}$) of C I ($1561 \text{ \AA}$), C II ($1336 \text{ \AA}$) will give a far more reliable estimate of the hydrogen/metal abundance ratio. It would be very interesting if different spiral arms gave significantly different results for this ratio because we might then perhaps be able to establish different age scales to stars in them. In connection with the mass distribution in the galaxy and with the heat balance of the interstellar gas an important unsolved problem

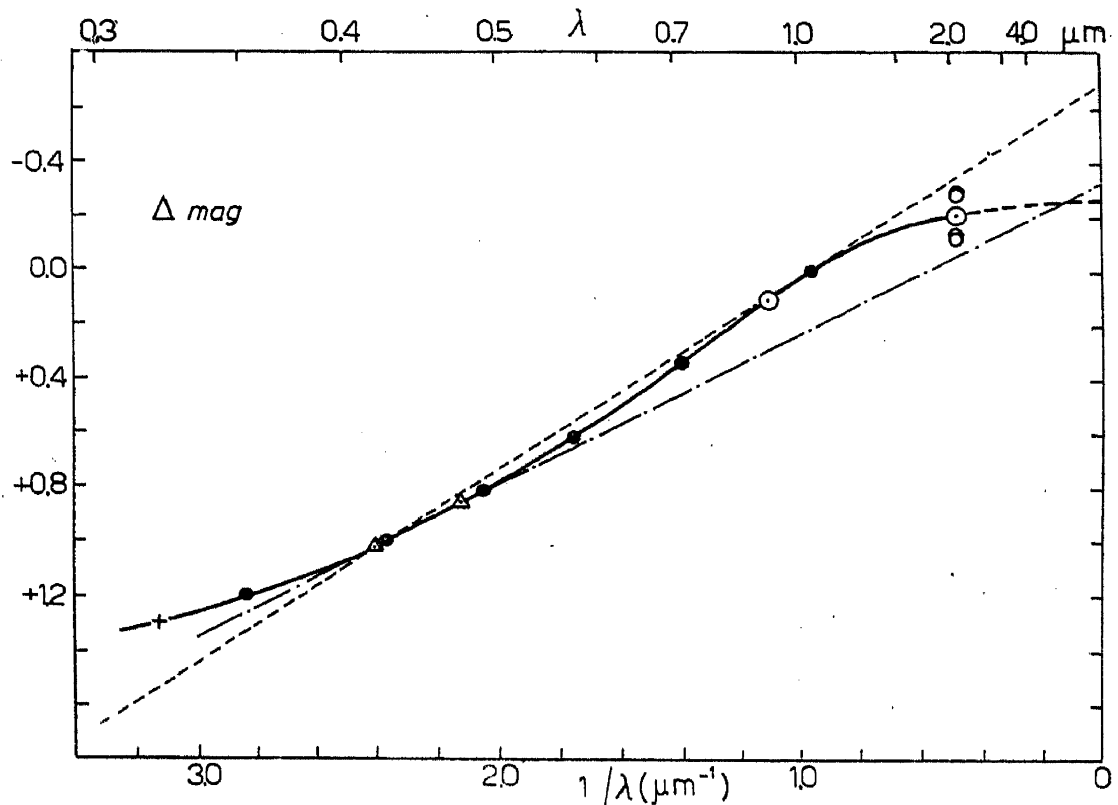


Fig. 7. - Normalized interstellar reddening curve.

is the question of the density of the H_2 molecule in appreciable abundance in interstellar space [7]. The measurement of the H_2 lines (1108 and $1000 \text{ \AA}$) would determine whether or not it is an important constituent of the interstellar matter.

Just as the Lyman continuum of hydrogen is expected to produce a considerable opacity so also will the interstellar dust although not to such a degree. Nevertheless any interpretations of the continuum radiation of stars will have to take into account the obscuration by the grains over the entire ultraviolet spectrum. While the effect of the grains is reasonably well known in the visible and near ultraviolet and is quite constant insofar as the wavelength dependence of the absorption is concerned, it is quite unlikely that extrapolations of the reddening curve will be very reliable. Figure 7 [8] shows the present extent of information on absorption by the interstellar dust. Fig-

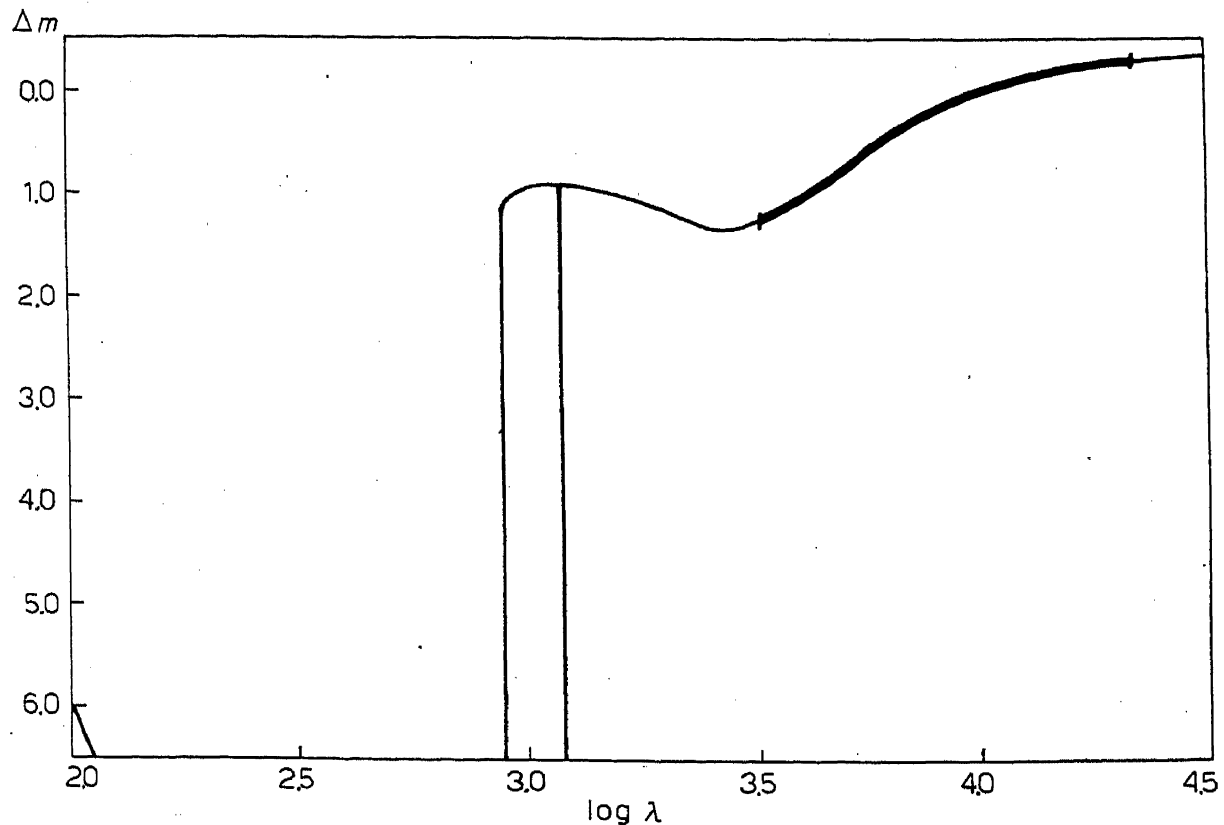


Fig. 8. - Transparency of interstellar matter of average density in the wavelength interval 100 to 31000 Å. The heavy segment at (3000 ÷ 20000) Å is the part that has been observed from the ground (see Fig. 7).

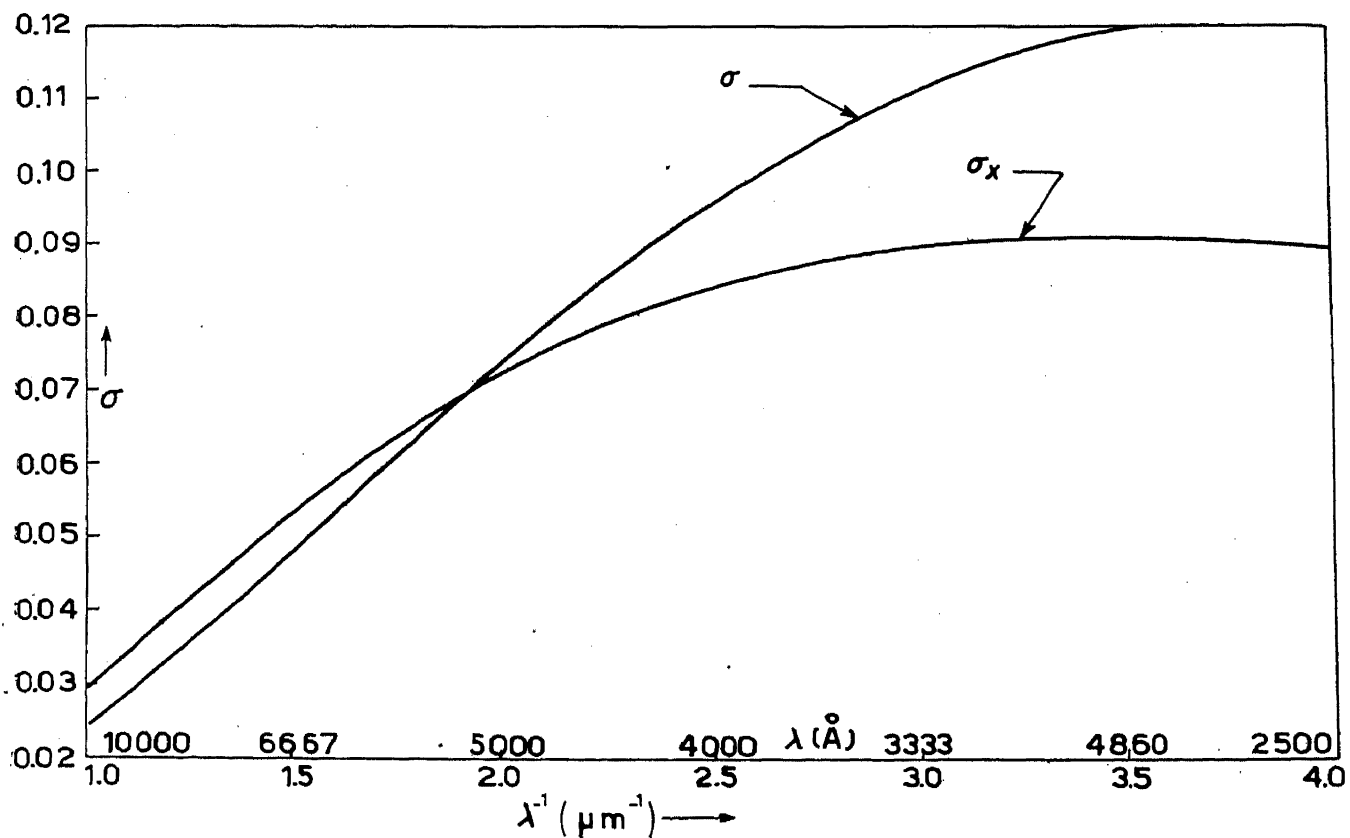


Fig. 9. - Wave length-dependence of extinction by a model dust cloud as viewed along and across magnetic field lines.

ure 8 [1] shows a possible extension of this curve into the far ultraviolet. The first vertical line is at Lyman α and the second one, where total obscuration is indicated, is at $\lambda = 912 \text{ \AA}$, the beginning of the Lyman continuum of hydrogen. However this extension could be easily in considerable error as can be seen in Fig. 9 [9] where it is shown that even for identical distributions of interstellar grains, the effect of orientation in varying degrees may alter the ultraviolet extinction by a considerable fraction. This effect has been calculated on the classical basis assuming the type of grains proposed by OORT and VAN DE HULST [10]. In this calculation the only property of grains which has been considered is the variation in their scattering with orientation. Naturally a more complete theory would involve such additional properties as the wave length-dependence of the index of refraction. It will be quite important, therefore, to extend the curve for interstellar absorption observationally, into an uncharted spectral region. Since the stellar continuum in this region is, as yet, unknown, and since this is usually a basic datum needed to determine the interstellar extinction one should attempt to by-pass this method, if possible. One such method has been proposed based on the emissions of diffuse nebulae. The relative intensities of the lines in the [O III] transition (involving the $^1S_0$, $^1D_2$, $^3P_{1,2}$ states) are well known from quantum mechanical data and therefore permit a determination of the extinction curve over a range from $\lambda 5007$ to $\lambda 2322 \text{ \AA}$. For example, all that need be known in comparing the observed intensities of $\lambda 4363$ and $\lambda 2322 \text{ \AA}$ are the corresponding f values. Since the longer wave length is well within the region of ground-based observations this would then connect the known extinction curve in an absolute way into the ultraviolet. Other such cases can be cited as perhaps especially useful. The converse of this problem is equally important. Much of the uncertainty regarding the temperature and composition of gaseous nebulae will be resolved upon adequately determining the wave length dependence of the extinction by the grains because this extinction alters the relative intensities of nebulae emission lines in different parts of the spectrum. It should be noted that, whereas the effect of the interstellar grains is commonly called interstellar reddening, it may well turn out that in the ultraviolet the effect may better be called interstellar blueing. The interstellar grains play an important role in such things as determination of galactic magnetic fields, temperatures of gas clouds and are a possible means by which interstellar molecular hydrogen is formed. Therefore it is crucial that observations be made which will help to decide which among the various theories of the grains is most nearly correct. The particular choice will have a considerable effect on general theories of the interstellar medium. It will also be related to the question of relative cosmic abundances of the elements although here the various theories of the grains do not differ by a large amount.

REFERENCES

- [1] A. D. CODE: *Astron. Journ.*, 65, 239 (1960).
- [2] H. E. HINTEREGGER: *Mém. Soc. Roi. Sci.*, vol. 4 (Liège, 1961), p. 111.
- [3] C. PECKER and F. ROERLICH: *Mém. Soc. Roi. Sci.*, vol. 4 (Liège, 1961), p. 265.
- [4] L. H. ALLER: *Pub. Astron. Soc. Pac.*, 71, 324 (1959).
- [5] L. GOLDBERG and E. R. DYER, Jr.: *Science in Space*, Chapter VII, *Galactic and extragalactic astronomy*, in *Nat. Acad. Sci.* (Washington, D.C., 1960). This chapter is a very valuable summary and introduction to the problems of astronomy in space.
- [6] A. BOGGESS: III, *Mém. Soc. Roi. Sci.*, vol. 4 (Liège, 1961), p. 459.
- [7] T. GOLD: *Mém. Soc. Roi. Sci.*, vol. 4 (Liège, 1961), p. 476.
- [8] A. E. WHITFORD: *Astron. Journ.*, 63, 201 (1958).
- [9] J. M. GREENBERG and A. S. MELTZER: *Ap. Journ.*, 132, 667 (1960).
- [10] J. H. OORT and H. C. VAN DE HULST: *B.A.N.*, 10, 187 (1946) (No. 376).

The History of Cosmic Radiation as Revealed by Isotopic Changes in the Meteorites and on the Earth.

J. GEISS, H. OESCHGER and U. SCHWARZ

Physikalisches Institut, University of Berne

1. - Introduction. (*)

In meteorites, in the atmospheres of the earth and at the surface of the moon nuclear transformations are continuously induced by cosmic radiation and solar high-energy particles at an appreciable rate. Detectable amounts of a number of stable and radioactive isotopes are produced, the determination of which gives information on a variety of subjects, such as the life history and origin of meteorites, the times scales of terrestrial, historical, pre-historical, and geological events, and the intensity, energy spectrum, spatial distribution and possible time variations of the cosmic and high-energy solar radiations.

In this paper we shall deal mainly with those aspects of the field which have bearing on cosmic radiation and solar radiation studies.

The meteorites, the lunar surface and the terrestrial atmosphere are far from being ideal targets for cosmic-ray experiments for the following reasons: The meteorites are objects of irregular shape, the preatmospheric size and shape of which is not known and may slowly change with time due to space erosion. Their orbital elements are not well known, and they are composed of different target nuclei. The spallation nuclei produced in the atmosphere are dispersed by very complex meteorological and geochemical processes into a variety of reservoirs, thus making a quantitative evaluation difficult.

However, the only information available to us on the history of the cosmic and solar high energy radiations is recorded by spallation products contained

(*) At the Varenna Summer School eight years ago F. G. HOUTERMANS [1] has given a lecture entitled *Cosmic Radiation in the Past*. It may be seen from that article, that many problems and methods of probing the history of cosmic radiation already were known or envisioned at that time. Many other problems or methods have arisen since. During the last eight years much more work has been done in this field, although we are still at the beginning only of understanding the record on the history of cosmic radiation which is preserved in nature.

in objects like meteorites or the layers of ice masses and deep sea sediments. It is mainly for this reason that almost any experimental effort to measure cosmic ray-produced isotopes and to elaborate theoretical attempts of interpretation are justified.

2. - The history of cosmic and solar radiations deduced from isotopes produced in the atmosphere.

The observed agreement between ^{14}C -ages and historical ages was the first information concerning the history of cosmic radiation (LIBBY [2]). The intensity was shown to have been constant over the last few thousand years within a few percent (Fig. 1). In addition, a reasonable agreement between

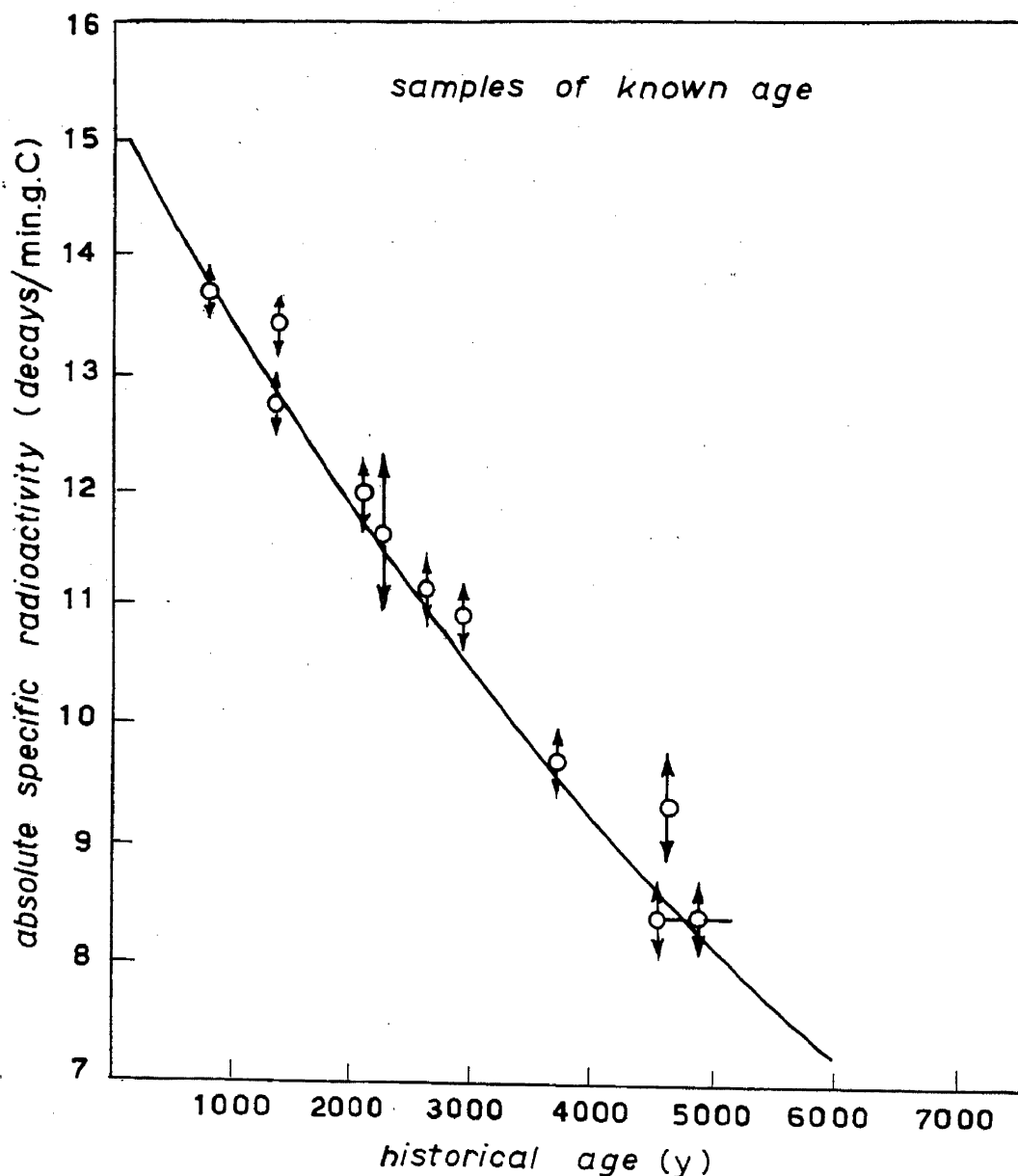


Fig. 1. - Predicted *vs.* observed radioactivities of samples of known age after LIBBY [3], showing that the cosmic radiation intensity has been essentially constant during the last several thousand years.

the ^{14}C and the Ionium/Protactinium time scales has been established for several 10^4 years (KULP and VOLCHOK [4]), ROSHOLT *et al.* [5]). As the errors of the Io/Pa method and similar methods are of the order of 6 000 years, the constancy of cosmic radiation over this period is established by this method only within a factor of two.

More recent accurate measurements on tree rings extending back over 3 000 years have shown time variations in the specific ^{14}C -activity of up to a few percent. However, these effects can not be interpreted unambiguously in terms of time variations of the cosmic ray intensity. They may be explained as well by a change of the rate of CO_2 -mixing between the atmosphere and other reservoirs, or by relatively small changes of the strength of the magnetic field of the earth. An apparent systematic change of the specific ^{14}C -activity might be simulated by an error in the adopted half life of 5 568 years. (DEVRIES [6]; WILLIS, TAUBER and MÜNNICH [7]; SUESS [8]).

The record on the history of high energy particle radiation over the last few million years is preserved by the ^{10}Be -concentrations in deep sea sediments (PETERS [9]; ARNOLD [10]). ^{10}Be is produced in the atmosphere and carried down by precipitation. The fraction which is precipitated into the sea is settling to the ocean floor in a relatively short time, probably within a few hundred years because of the low solubility of beryllium. The ages of different layers of the sedimentary column can be obtained from radioactive methods (protactinium-ionium, potassium-argon), and sedimentation rates may be calculated. From this the ^{10}Be production rate $P(t)$ per cm^2 sediment can be inferred as a function of time by the relation:

$$(1) \quad P(t) = R \cdot [^{10}\text{Be}] \cdot \exp [\lambda t],$$

where R is the rate of sedimentation in grams per cm^2 and year, $[^{10}\text{Be}]$ is the concentration of ^{10}Be per gram, and t the age of the sediment layer under investigation.

$P(t)$ is related to the total cosmic and solar radiation flux. ^{10}Be measurements and their interpretation and the dating of deep sea sediments are very difficult problems, but in principle this method would enable us to look for intensity variations during the last several million years with a time resolution of a few thousand years. Thus, it is suitable for detecting short high-intensity «bursts» of high-energy radiation. Events such as solar bursts of a far greater magnitude than observed during the short period of historical observation, or novae or supernovae explosions in the vicinity of the solar system are detectable by this method.

Variations of the cosmic radiation intensity during the last 50 to 80 years may be detected by measuring the ^3H -contents in the layers of the ice of glaciers and ice caps (BEGEMANN [11]). The dating in this case is done by

observing and counting the seasonal variations in the crystalline structure of the ice or in the isotopic variations of oxygen or hydrogen (cf. EPSTEIN [12]). Such measurements would give an integral picture of the intensity variations connected with the solar cycle. A positive correlation with the solar activity would mean that during the active period the solar output more than compensates the decrease of the galactic flux caused by the ejection of plasma from the sun, whereas a negative correlation would mean the opposite. First measurements by BEGEMANN [11] indicate a negative correlation. Comparative measurements of ^3H -variations in polar ice caps and in glaciers at lower latitude might reveal something on the energy of the particles which are affected by these changes.

The measured inventory of ^{14}C on earth agrees very well with the production rate calculated from the flux of galactic cosmic radiation. This agreement places an upper limit on the total output of solar high-energy particles (LAL and PETERS [13]).

The spallation nuclei produced in the atmosphere are not only being used as radiation monitors, but also applied for the investigation of fundamental geophysical problems or historical questions. Often methods based on spallation isotopes have given answers to geophysical problems of a far more general character than had been obtained before with classical methods. So, the general circulation in the atmosphere has been studied, exchange times between the troposphere and stratosphere or the northern and southern hemispheres have been determined and air masses can be followed along their way from the stratosphere to the troposphere (cf. LAL and PETERS [13]). The complex exchange system between the atmosphere, the ocean and the biosphere has been studied and exchange rates have been ascertained (CRAIG [14]). Also it was possible to directly measure the velocity of deep sea currents (RAKESTRAW *et al.* [15]; BIEN *et al.* [16]) and to develop models for the mixing of oceanic water masses (BROECKER *et al.* [17]).

3. - Spallation isotopes in meteorites.

It is known from radioactive dating that meteorites are very old, many of them having ages of more than 4 billion years (*). These radioactive ages refer to the time at which the meteorite attained its chemical composition and its crystal structure. However, nothing is inferred about the length of time during which the meteorite has been exposed to the cosmic radiation. In fact, all theories of meteorite formation (cf. UREY [20]; ANDERS and

(*) For further studies of the general subject of this chapter see ZÄHRINGER [18] in this volume and EBERHARDT and GEISS [19] and references therein.

GOLES [21]) suppose that the meteorites originally were formed as parts of much larger bodies and were separated from these at some later date, presumably by collision processes.

It is generally believed that meteorites originate in the asteroidal belt, although very little direct proof for this does exist. A precise calculation of an orbit from photographs obtained with rotating-shutter cameras has been possible only in the case of the chondritic meteorite Pribram, which fell in 1959, leading to a perihelion distance of 0.8 a.u., and an aphelion distance of 4 a.u. (CEPLECHA [22]); (cf. Fig. 2).

Estimates on the orbits of other meteorites which are based on visual observations by untrained observers are of very poor accuracy, but usually have led to similar orbital elements.

In this discussion we shall assume therefore, that the meteorites generally have aphelion distances of (3 ÷ 5) a.u. and perihelion distances of 1 a.u. or less, although there may quite possibly exist some meteorites with essentially different orbits. Even the possibility that some meteorites come from the moon cannot be excluded.

After the meteorite has been liberated from its parent body, its matter is exposed to cosmic radiation, and many stable and radioactive spallation products are formed. The irradiation is terminated at the time of fall of the meteorite, because the magnetic field and the atmosphere of the earth cut down the level of irradiation by some orders of magnitude.

Cosmic ray-produced nuclei in most cases can be readily identified. Radioactive isotopes with $T < 10^8$ years cannot have survived since the cessation of the nucleosynthesis processes and therefore must have been formed by irradiation more recently (intermediate products of the three radioactive series are excluded). Stable and long lived spallation products are recognized from abnormal isotopic compositions. Their detectability is restricted to rare isotopes of rare elements, *e.g.* most of the rare gas isotopes, ^{40}K , ^{45}Sc , etc. Further proof for the cosmic-ray origin of these isotopes are the dependence of their concentrations on the size of the meteorite (BAUER [23]) and the depth-dependence within the meteorite itself (HOFMAN and NIER [24]; SIGNER and NIER [25]).

The length of time during which a meteorite was exposed to cosmic radiation may be calculated under certain assumptions. We define the «radiation

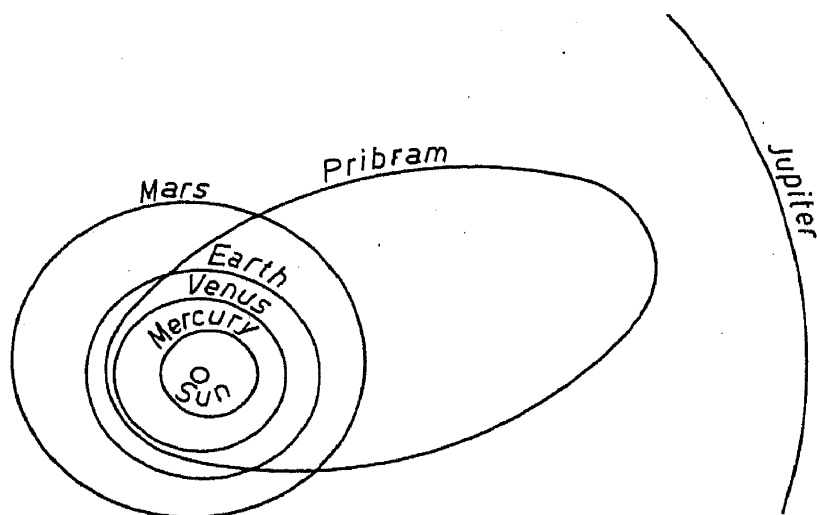


Fig. 2. - Orbit of the chondritic meteorite Pribram calculated by CEPLECHA [22].

age » of a meteorite by the following relation:

$$(2) \quad T_s = \frac{P_R}{P_{ST}} \cdot \frac{N_{ST}}{dN_R/dt},$$

where N_{ST} is the number of atoms of a stable spallation product, dN_R/dt the number of decaying atoms per unit time of a radioactive spallation product, and P_R and P_{ST} are the respective production rates. T_s gives the time of the separation of the meteorite from a larger parent body if

a) the meteorite was incorporated in the parent body in such a way that it was completely shielded from cosmic radiation,

b) the meteorite was not changed in size or shape by further collisions or by space erosion after the separation from the parent body,

c) the cosmic ray intensity, composition and spectrum was constant during the time of exposure,

d) P_R/P_{ST} can be estimated with sufficient accuracy.

P_R/P_{ST} is obtained from cross-section measurements with artificially accelerated protons and α -particles at various energies. In a very small meteorite the P -values would be obtained by integrating the cross-sections over the cosmic ray spectrum. In a meteorite of finite size, the energy spectrum is modified because of absorption, slowing down and secondary production, but it can be deduced from the concentrations of several spallation nuclei (cf. Section 4) and meaningful radiation ages may be obtained.

If such pairs of a stable and a radioactive isotope are taken, for which the ratio of their production cross-sections does not depend greatly on the energy, radiation ages may be calculated safely from a pair alone. This is the case if the two nuclei are neighbours far away from the target nucleus; *i.e.* the pairs $^{38}\text{A}/^{39}\text{A}$ or $^{36}\text{A}/^{36}\text{Cl}$ are well suited for the determination of radiation ages in iron meteorites.

In stone meteorites, radiation ages calculated from the $^3\text{He}/^3\text{H}$ ratio are relatively reliable, because their production rates depend much less on the chemical composition than the production rates of remnant nuclei.

So far radiation ages have revealed some very interesting features:

1) In all cases the radiation ages are much smaller than the radioactive ages.

2) Iron meteorites and stone meteorites have very different radiation ages. Most of the irons give $(10^8 \div 10^9)$ years, whereas most of the stones give values smaller than 10^8 years.

3) There appears to be a grouping of radiation ages if certain meteorite classes are regarded separately. The $^3\text{He}/^3\text{H}$ radiation ages of chondrites show such a concentration at about $22 \cdot 10^6$ y. (GEISS, OESCHGER and SIGNER [26]) suggesting a definite collision process at that time, which created many of the chondritic meteorites (Fig. 3). Recently WÄKNE [28] and HEYMANN and SCHAEFFER [29] have suggested, that also for iron meteorites there is a grouping of radiation ages, in this case at about 500 to 600 million years.

The obvious explanation for the striking discrepancy between the radiation ages of iron and stone meteorites would seem to be space erosion, because of their different crushing strengths (FIREMAN and DE FELICE [30]). In case of a constant erosion rate γ and constant cosmic ray intensity the radiation age T_s , as defined in eq. (2), measured in the center of a spherical meteorite, is connected with the break-up time T_b by the following equation (WHIPPLE and FIREMAN [31]).

$$(3) \quad T_s = \frac{L}{\gamma} \left\{ 1 - \exp \left[-\frac{\gamma}{L} T_b \right] \right\},$$

(L is a characteristic absorption length for the cosmic radiation).

Two limiting cases may be distinguished:

$$\gamma T_s \ll L \quad (\text{low erosion rate}): \quad T_s = T_b,$$

$$\gamma T_s \gg L \quad (\text{high erosion rate}): \quad T_s = L/\gamma.$$

The relatively small crushing strength of stone meteorites therefore seemingly would explain their comparatively low radiation ages, which would be given essentially by the erosion rate.

Several arguments, however, may be put forward against this. One is, that not the macroscopic crushing strength but the hardness of the individual grains in stone meteorites determines the rate of erosion caused by dust (GEISS and OESCHGER [32]). Another argument, put forward by UREY [33], is that strong space erosion would lead to a size distribution of stone meteorites which disagrees with the observation. Another one is, that there exist stone meteorites having radiation ages greater than 10^8 years (BEGEMANN, GEISS and HESS [34]).

We believe, therefore, that at least those radiation ages of meteorites which are of the order of 10^7 years are determined by one or several collision processes between sizable objects and are very little influenced by erosion. This would be substantiated by the finding of the mentioned grouping in the radiation ages. However, we do not want to exclude the possibility that the ab-

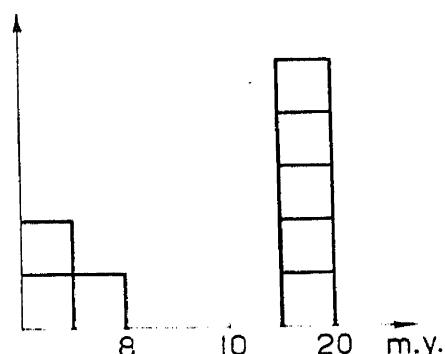


Fig. 3. - Radiation ages of chondrites as measured by the $^3\text{He}/^3\text{H}$ -method after GEISS, OESCHGER and SIGNER [26], including values of GOEBEL *et al.* [27], showing a grouping around $22 \cdot 10^6$ y.

sence of radiation ages of stone meteorites of several hundred million years and more might be due to some preferential erosion mechanism.

There actually exists a very direct way of deciding the question whether space erosion is a determining factor in the radiation ages of stone meteorites or not. ^{36}A is produced in appreciable amounts in meteorites by slow neutrons from $^{35}\text{Cl}(n, \gamma)^{36}\text{Cl}$, as has been calculated recently by EBERHARDT, GEISS and LUTZ [35]. The flux of slow neutrons is strongly size-dependent so that an appreciable production of ^{36}A by this process is only possible in meteorites above a certain size (about 30 cm radius, cf. Fig. 4). If necessary, this ^{36}A may

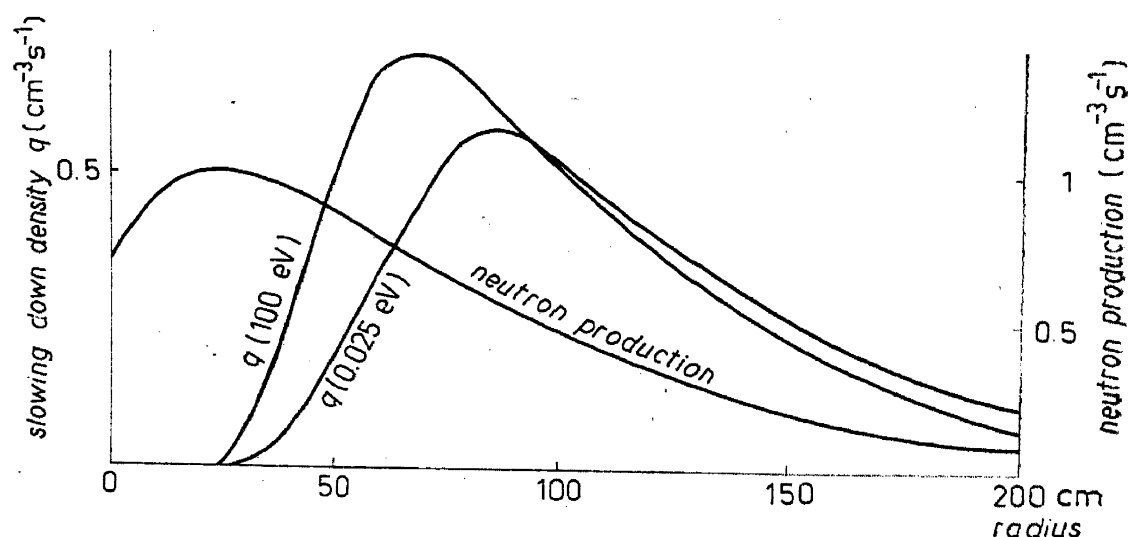


Fig. 4. - Slowing down density q of neutrons at two different Fermi-ages in center of chondritic meteorite. Values given are without absorption. Actual slowing down density is obtained by multiplying q with the resonance escape probability (after EBERHARDT, GEISS and LUTZ [35]).

be distinguished from ^{36}A of other origin by looking into chlorine-containing components in the meteorite. If then large and very small stone meteorites contain this neutron-produced ^{36}A in comparable concentrations relative to chlorine, the small meteorites must once have gone through the optimal size region for slow neutron fluxes which lies roughly between 50 and 200 cm radius and space erosion would indeed be the factor determining the life-time of a stone meteorite in space.

The oldest radiation ages of iron meteorites found so far are about $(1 \div 2)$ billion years (SPRENKEL [36]; FISHER and SCHAEFFER [37]). This gives a lower limit for the time-integral of the cosmic radiation intensity, which is

$$\int I dt = I_0 T_s^{\max}.$$

An upper limit for the dust density in space along the path of the iron meteorites, averaged over the last one or two billion years can be derived from

the consideration that some of these meteorites have survived dust bombardment for this length of time. A value of the order of 10^{-22} g/cm³ is obtained (cf. GEISS and OESCHGER [32]). This upper limit is considerably lower than dust densities derived from observations on the zodiacal light (VAN DE HULST [38]; ELSÄSSER [39]), but of course the two methods are determining dust densities at quite different regions in space.

The presence of cosmic ray-produced isotopes in objects of questionable origin can be used as definite proof for an extraterrestrial source. This argument has been of considerable importance in the discussion of the origin of tektites (ANDERS [40]).

By measuring the concentration of cosmic ray-produced isotopes in lunar material it will be possible to investigate the history of the surface of the moon. The radiation ages of lunar samples will help to decide whether the lunar surface is very ancient or whether it is undergoing changes through differentiation processes (O'KEEFE [41]) or through dust movement (GOLD [42]), and will allow quantitative estimates of the time scales involved.

4. - Spallation reactions.

In order to quantitatively compare the observed abundances of spallation nuclei with their expected production rates we have to discuss the properties of spallation reactions.

Accelerator experiments have shown that the cross-sections of the various remnant nuclei produced by spallation from a target nucleus of mass A_0 are smooth functions of the kinetic energy E and of the total mass loss ΔA (RUDSTAM [43]; BARR [44]). The total cross-section of an isobar (sum of the cross-sections of all nuclei of an isobar) can be represented by the following relation (cf. HONDA and LAL [45]):

$$(4) \quad \sigma(\Delta A, E) = \sigma_0 A_0^{\frac{2}{3}} c E^{-\frac{2}{3}} \exp [-c E^{-\frac{2}{3}} \Delta A]$$

an extension of the formula given by RUDSTAM [43]. This formula is experimentally verified for protons between 50 MeV and 5.7 GeV. Some experimental evidence exists that there is little difference, if any at all, between the production rates of protons, α -particles and neutrons. For light spallation products the main contribution to the production is from the directly produced or evaporated fragments. This effect is negligible for spallation products with mass $A > 20$, at least up to 5.7 GeV. For $\Delta A = 1$ to 5, deviations from eq. (4) occur; for $\Delta A = 2$ to 5, these deviations are restricted to higher energies.

In eq. (4), $\sigma(\Delta A, E)$ is normalized so that

$$\int_0^{A_0} \sigma(\Delta A, E) d(\Delta A) = \sigma_0 A_0^{\frac{2}{3}}$$

gives the total inelastic cross-section which has been shown experimentally to be approximately independent of the energy and proportional to the geometric cross-section. c is a constant depending only on A_0 . This A_0 -dependence which is derived by OESCHGER and SCHWARZ [46] is important for the interpretation of the abundances of cosmic ray-produced isotopes in stone meteorites. The iron-cobalt-nickel alloy of most iron meteorites can be represented with sufficient accuracy by the nucleus ^{56}Fe . For this target nucleus the constant c is $0.11 (\text{GeV})^{\frac{2}{3}}$.

If $f(E)dE$ is the differential (kinetic) energy spectrum of irradiation, the production rates are calculated by the equation:

$$(5) \quad N(\Delta A) = \int_0^{\infty} f(E) \sigma(\Delta A, E) dE,$$

$N(\Delta A)$ is the production rate of an isobar per iron atom.

Recently, in an extensive study, ARNOLD, HONDA and LAL [47] have calculated by numerical integration the production rates of the various spallation products at different depths in iron meteorites on the basis of the observed cosmic ray spectrum, the development of secondaries in the atmosphere and spallation excitation functions. By comparison with measured concentrations of radioactive spallation nuclei in the iron meteorite Aroos (cf. Section 6), they conclude that all these nuclei are in equilibrium, indicating an essentially constant cosmic ray flux.

In this paper we shall use a somewhat different approach. We shall attempt to keep an analytical treatment for eq. (5) throughout the discussion. $N(\Delta A)$ is an integral transform of $f(E)$ which transforms the energy spectrum of irradiation into the ΔA spectrum. In fact, if

$$(6) \quad s = cE^{-\frac{2}{3}}$$

is substituted, we obtain

$$(7) \quad N(\Delta A) = \int_0^{\infty} g(s) \exp[-\Delta A s] ds,$$

with

$$(7a) \quad g(s) = k_0 c^{\frac{3}{2}} f \left[\left(\frac{c}{s} \right)^{\frac{3}{2}} \right] s^{-\frac{3}{2}}, \quad k_0 = \frac{3}{2} \sigma_0 A_0^{\frac{3}{2}},$$

where $N(\Delta A)$ is the Laplace-transform of $g(s)$. This enables us to explicitly calculate the energy spectrum of irradiation through the inverse transform:

$$(8) \quad f(E) dE = \frac{1}{2\pi i} \frac{dE}{k_0 E} \int_{\alpha-i\infty}^{\alpha+i\infty} \exp [cE^{-\frac{3}{2}} \Delta A] N(\Delta A) d(\Delta A),$$

with $\Delta A = x + iy$.

For instance

$$(9) \quad f(E) dE = f_0 E^{-\alpha} dE$$

results in a spallation product distribution

$$(10) \quad N(\Delta A) = \frac{f_0 k_0 c^{-(n-1)}}{(\Delta A)^n} \Gamma(n),$$

with

$$(11) \quad \Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx$$

and

$$(12) \quad n = \frac{1}{2}(3\alpha - 1).$$

In this treatment we have always extended the integration from 0 to ∞ , although at these limits neither the energy spectrum nor the excitation function $\sigma(E, \Delta A)$ in the form used here are consistent with the experimental facts. It is shown in the following discussion, that these limits of integration are of little influence in the result.

The analytical treatment used here enables us to calculate very easily the energy range mainly responsible for the production of a certain isobar ΔA . For an illustration we take $\alpha = 1.67$ and obtain $N(\Delta A) \sim 1/(\Delta A)^2$, an α -value which roughly corresponds to effective energy spectra in meteorites (*). The contribution of the different energies of the production to some ΔA isobars is plotted in Fig. 5. The energy of maximum production E_m is given by

$$(13) \quad E_m = 5.7 (\Delta A)^{\frac{2}{3}} [\text{MeV}]$$

(*) An integer n is chosen because the following relations (13) and (14) can be expressed in simple terms.

and the median energy E_p (*) by

$$(14) \quad E_p = 17.3(\Delta A)^{\frac{1}{3}} \text{ [MeV]}.$$

The energy range between E_m and $2E_p$ is responsible for 60 % of the production for a given ΔA isobar (Table I). Thus, for all practical purposes, there exists

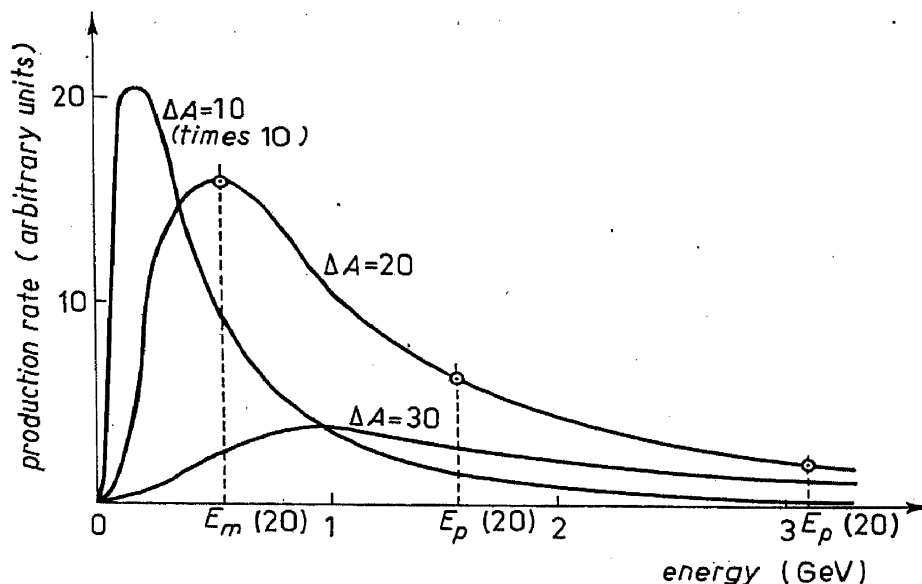


Fig. 5. - Contribution of different energies to the production of certain isobars. An irradiation spectrum $f(E)dE = f_0 E^{-1.67} dE$ is assumed (E = kinetic energy). The energy range between E_m and $2E_p$ produces 60 percent of the isobar (cf. Table I).

a characteristic limited energy range which is responsible for the production of an isobar. If the range of considered ΔA 's is restricted, only a limited energy range contributes appreciably to the production. Thus, the extreme energy ranges ($\rightarrow 0$ and $\rightarrow \infty$) are of little influence on the ΔA distribution. On the other side, the spectral form in these extreme ranges can not be inferred from the limited ΔA range. From this follows that in practical cases the integral in eqs. (5) and (6) may be safely extended from 0 to ∞ , if only the spectral form is known over the range which appreciably contributes to the production.

5. - The application to stable isotopes in meteorites.

In an infinitesimally small iron meteorite the ΔA distribution directly leads to the energy spectrum of the cosmic radiation. In a practical case the pos-

(*) I.e., 50% of production above and 50% below E_p .

sible contamination of the iron-cobalt-nickel alloy with elements of lower atomic number has to be taken into account. For instance sulfur, contained in troilite inclusions has a production cross-section for ^{21}Ne which is 7 times higher than that for iron.

For a meteorite of finite size the original energy and mass spectrum of the cosmic or solar radiation is modified by ionization, absorption and the production of secondaries, including mesons. Here the ΔA distribution represents the actual irradiation spectrum at the site of the sample.

Because of the long irradiation times involved the stable spallation isotopes in meteorites in most cases represent the total isobar. Therefore, they directly give a $N(\Delta A)$ distribution. In Fig. 6 average rare gas spallation isotope abundances in the iron meteorite Grant (SIGNER and NIER [25]) are plotted. A ^{45}Sc value is included obtained from average $^{45}\text{Sc}/^{22}\text{Ne}$ ratios of relatively small iron meteorites given by WÄNKE [48]. ^{38}A , which does not represent the total isobar, has been corrected according to the cross-sections given by BARR [44]. A distribution

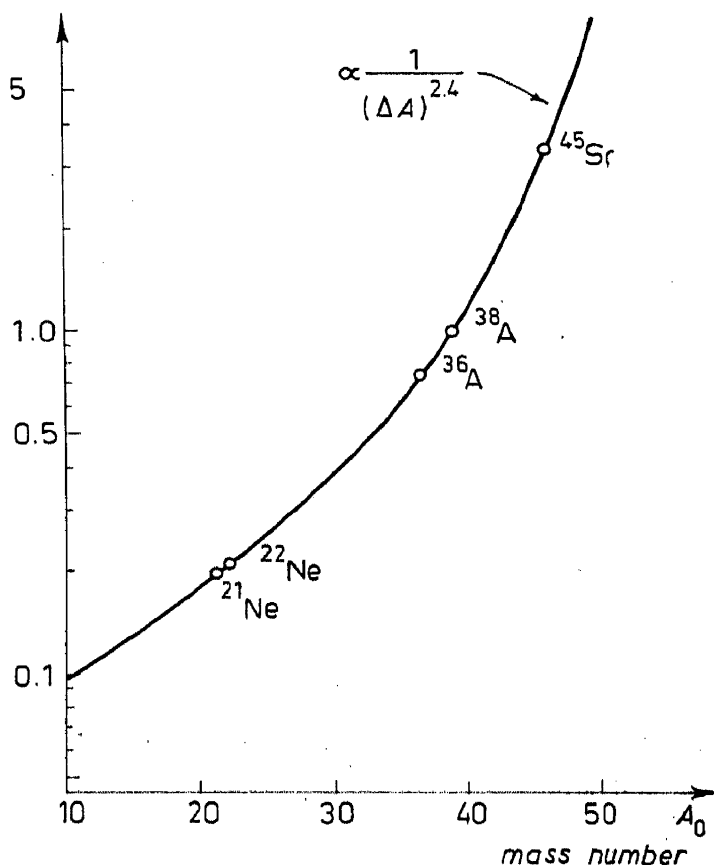


Fig. 6. - Abundances of cosmic ray-produced isotopes ^{21}Ne , ^{22}Ne , ^{36}A and ^{38}A measured by SIGNER and NIER [25] in the Grant meteorite. A ratio $^{38}\text{A}/^{21}\text{Ne} = 5.0$ is taken. The ^{36}A -value is corrected after BARR [44]. ^{45}Sc is included from the average ratio $^{45}\text{Sc}/^{22}\text{Ne} = 18$ found by WÄNKE [48] in iron meteorites smaller than 500 kg. Curve gives $(\Delta A)^{-2.4}$ distribution.

$$(15) \quad N(\Delta A) = \frac{a_s}{(\Delta A)^{2.4}},$$

gives a very good fit with the experimental values (cf. Fig. 6), corresponding to an energy spectrum

$$(16) \quad f(E) dE = \frac{f_0}{E^{1.93}} dE.$$

Recently HONDA and STAUFFER [49] have succeeded in measuring spallation-produced ^{43}Ca and ^{50}V . The inclusion of these isotopes into a plot like Fig. 6

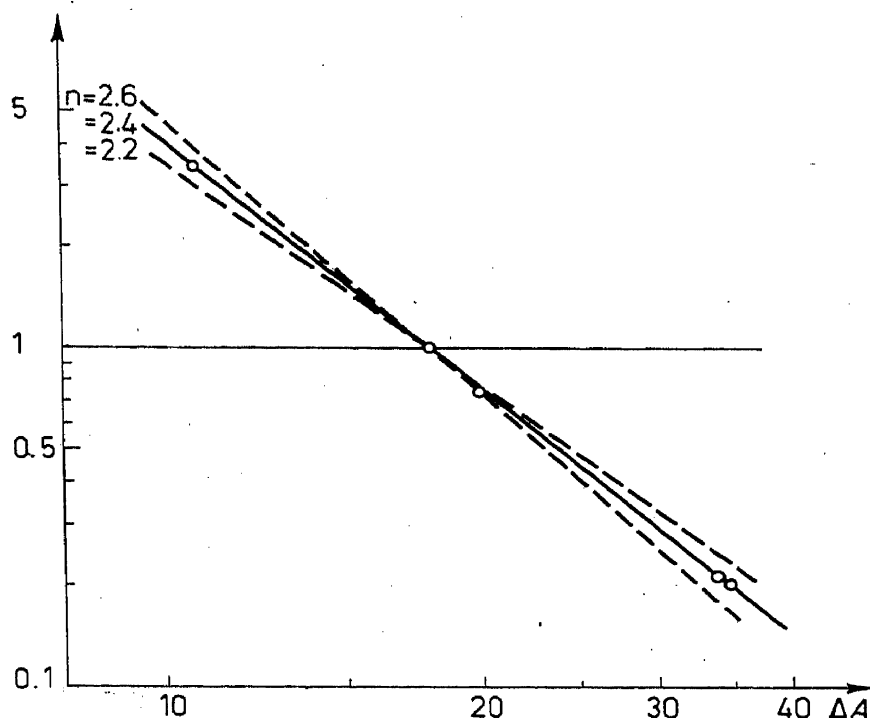


Fig. 7. — Log N vs. log ΔA plot of stable spallation nuclei in the Grant meteorite (cf. Fig. 6). Dashed lines corresponding to energy spectra with $\alpha = 1.80$ and $\alpha = 2.07$ indicate how sensitively ΔA distribution reacts to change in irradiation spectrum.

will extend the ΔA range appreciably. From Fig. 7 it can be seen how sensitive a $N(\Delta A)$ vs. ΔA plot is in determining irradiation spectra.

6. — The intensity of cosmic radiation during the last 10^7 years.

The comparison of spallation nuclei of different half-lives in meteorites affords an insight into the history of the cosmic and solar radiations. In principle, the method is to compare the concentrations of the radioactive nuclei with their production cross-sections. From this it can be judged whether the nuclei of different half-lives are in equilibrium or not, which is a check on the time-constancy of the radiation intensity. It was shown in Section 3 that many meteorites have radiation ages much longer than 10^7 years and that space erosion for periods of 10^7 years or less can probably be neglected for stone meteorites and certainly for irons. So meteorites are workable monitors of the radiation intensity during the last 10^7 years.

Two qualifications, however, have to be made. In the first place, it can not be excluded that one or the other meteorite has suffered a chance collision during the last 10^7 years. This would result in a change of the degree of shielding for a particular piece of the meteorite and would simulate an increase of the radiation intensity with time. It can be assumed that this has happened only to a small fraction of the meteorites so that this problem may be

overcome by investigating several different meteorites. Secondly, during the last 10^7 years the orbital elements of a meteorite may have appreciably changed due to perturbations by the planets. Therefore, it should be born in mind that during this period the meteorite perhaps is scanning different parts of the interplanetary space with different radiation intensities at different times, and time variations may be simulated by this effect.

HONDA, SHEDLOVSKY and ARNOLD [50] have measured for the first time a large number of radioactive spallation nuclei in the same sample of an iron meteorite, the meteorite Aroos which fell in 1959. In addition, FIREMAN and DE-FELICE [30] have measured the ^{37}A and ^{39}A concentrations in the same meteorite. Already in their first publication HONDA, SHEDLOVSKY and ARNOLD [50] point out that all these isotopes with different half-lives have abundances which indicate equilibrium within the limits of experimental error and the uncertainty of estimated production rates. This was substantiated by the work of ARNOLD, HONDA and LAL [47], as was mentioned before.

For the discussion we calculate from the measured activities of the radioactive isotopes the production rates of the total isobars in the meteorite sample, taking the relative cross-sections for the nuclei of an isobar from the experimental and interpolated figures given by BARR [44] (*). These values, called «activities of isobars», are plotted *vs.* ΔA in Fig. 8. A distribution in agreement with the $(\Delta A)^{-2.4}$ distribution of the stable isotopes in the Grant meteorite is obtained, indicating a very similar irradiation spectrum in the two samples. In Fig. 9 the «activities of isobars» are multiplied by $(\Delta A)^{2.4}$ and plotted *vs.* the half-life. In this way the experimental activities of the isotopes with different half-lives are normalized to their expected production rates so that a systematic trend with time would appear. This normalization is based on the observed ΔA distribution in this meteorite (Fig. 8) and the observed

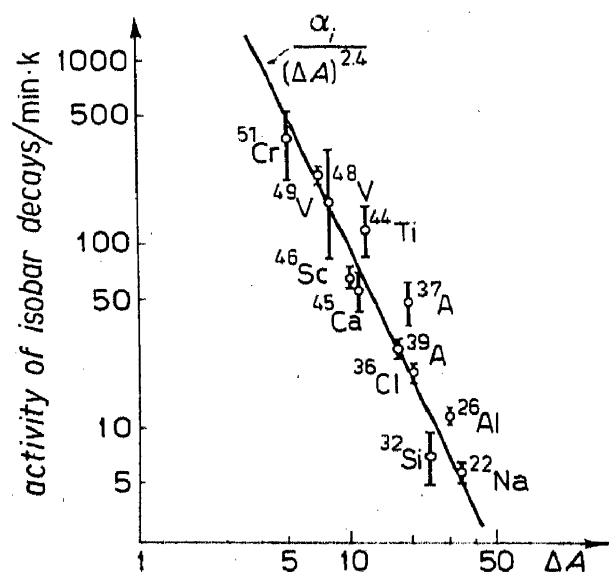


Fig. 8. — Radioactive spallation products in the iron meteorite Aroos measured by HONDA, SHEDLOVSKY and ARNOLD [50], including ^{37}A and ^{39}A values by FIREMAN and DE FELICE [30]. The «activities of the isobars» derived from the measured activities of spallation nuclei are plotted *vs.* ΔA .

(*) Barr's experiments were performed at 5.7 GeV with a copper target. These data may be taken as average values because the distribution of spallation nuclei is practically independent of the energy (RUDSTAM [43], BARR [44]).

ΔA distribution of the stable isotopes in a meteorite of not too different size (Fig. 6 and 7). The second argument would be much more straightforward if the stable isotope distribution would be taken from the same meteorite. Unfortunately stable isotope data of Aroos are not yet available to us (*).

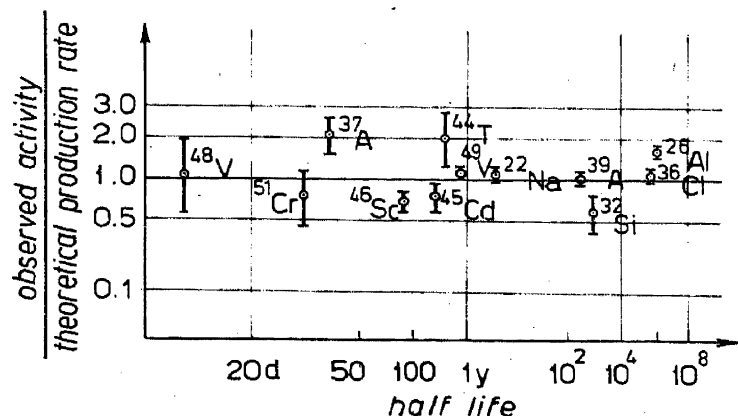


Fig. 9. — Ratio of observed activity to theoretical production rate of spallation products in Aroos iron meteorite vs. half-life.

The measured activities of all the radioactive isotopes agree within a factor of two with the theoretical production rates, and no systematic trends can be discerned. As our calculations do not take into account fragmentation products, Fig. 8 and 9 do not include ^{10}Be . Also ^{53}Mn , with $\Delta A = 3$, is not included, as no stable isotopes with low ΔA were measured, which would control the irradiation spectrum at the low energies. The calculations of ARNOLD, HONDA and LAL [47]

indicate that also these isotopes are in equilibrium. The measured ^{26}Al concentration is relatively high. This could be due to a small amount of troilite in the sample (ARNOLD, HONDA and LAL [47]). The relatively high abundance of ^{37}A will be discussed in Section 8.

Within the limits of error the radioactive spallation products with half-lives from 15 days to $2.5 \cdot 10^6$ years or more (^{53}Mn) seem to be in equilibrium. This shows that the cosmic plus solar radiation intensities have been essentially constant during the last several million years, within the limits discussed in the next Section.

7. — Detectability of intensity variations.

If $I(t)$ is the radiation intensity as a function of time, the number $N(\lambda)$ of nuclei of a radioactive isotope with a decay constant λ present immediately after the fall of the meteorite is given by the Laplace transform

$$(17) \quad N(\lambda) = \sigma_{\lambda} \int_0^{\infty} I(t) \exp[-\lambda t] dt,$$

(*) If the stable and the radioactive spallation-product distributions in a meteorite sample yield the same exponent n (cf. eq. (10)), a constancy of the irradiation spectrum at the site of the sample during the period of exposure of the meteorite is indicated.

where σ_λ is an average cross-section for the production of the isotope with decay constant λ . $N(\lambda)$ is given in numbers of radioactive nuclei per target nucleus.

The irradiation intensity, as a function of time, can be calculated explicitly by means of the inverse Laplace-transform (cf. eq. (8)) from the measured abundances $N(\lambda)$ of radioactive spallation products with different decay constants. Of course, a sufficiently accurate continuous function $N(\lambda)$ can be given only if enough different spallation products are precisely measured.

For constant radiation intensity, $N(\lambda)$ becomes

$$(18) \quad N(\lambda) = \sigma_\lambda \frac{I_0}{\lambda},$$

$N(\lambda)$ is in equilibrium with the radiation, *i.e.* production and decay rates are equal.

In the following we consider two types of possible time variations of the radiation intensity and their transformation into the « λ -space ». These are the harmonic oscillation, and the short burst.

From

$$(19) \quad I(t) = I_0 + I_1 \sin(\omega t - \varphi)$$

one obtains

$$(20) \quad \frac{\lambda}{\sigma_\lambda} N(\lambda) = I_0 + I_1 \frac{\lambda}{\sqrt{\lambda^2 + \omega^2}} \sin(\alpha - \varphi), \quad \text{with } \operatorname{tg} \alpha = \frac{\omega}{\lambda},$$

$N(\lambda)$ is a function of λ/ω and of φ . φ is the phase of the harmonic component of the radiation at the time of fall.

The periodicity of the radiation intensity results in a periodicity of $N(\lambda)$ with respect to the time of fall. However, the periodic component of the radiation is by a factor $\lambda/\sqrt{\lambda^2 + \omega^2}$ less effective than the constant component, which has to be born in mind when expressing the content of Fig. 9 in terms of intensity variations. An instructive example is the reaction of the ^3H -concentration to irradiation fluctuations with the solar cycle. Here, the factor $\lambda/\sqrt{\lambda^2 + \omega^2}$ is 0.096. Therefore, tritium has too long a half life to pick up efficiently presumed solar-cycle variation in meteorites. Its concentration should depend relatively little on the time of fall of the meteorites. Thus tritium is sufficiently averaging intensity variations resulting from the 11 years cycle and may be safely used for the calculation of radiation ages of meteorites.

^{22}Na is a suitable isotope for detecting solar-cycle intensity variations in meteorites, because it is averaging sufficiently well over the orbit, but not over the 11 years cycle.

If we assume a short radiation burst at time t_0 the dosage $\int I dt$ of which may equal $I_0 t_0$ (the total dosage the meteorite has received after this event), we obtain

$$(21) \quad \frac{\lambda}{\sigma_\lambda} N(\lambda) = I_0 \{ 1 + \lambda t_0 \exp[-\lambda t_0] \}.$$

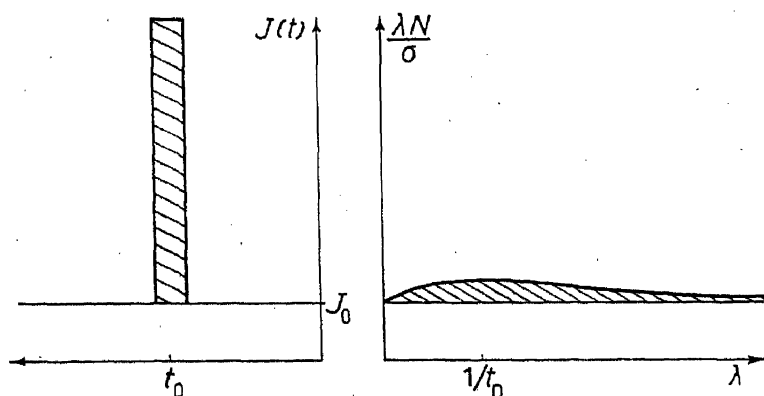


Fig. 10. — Influence of cosmic ray burst (shaded area on left) on abundance of radioactive spallation nuclei with decay constant λ (shaded area on right).

Even for an isotope with $\lambda = 1/t_0$, which is the optimum for the detection of this burst, $N(\lambda)$ increases only by 36 % (Fig. 10). Such an increase would not be noticeable in Fig. 9.

Thus meteorites are comparatively insensitive detectors for bursts. In the earth atmosphere, where spallation products such as ^{10}Be are continuously removed from the irradiation zone and incorporated

in stratified layers, the situation is much more favorable, as a differential time record is preserved. In principle here a burst with $\int I dt = I_0 \cdot 10^4$ years is detectable even if it occurred a few million years ago.

8. — Comparison of radiation intensities at 1 a.u. and 3 to 4 a.u.

Isotopes with half-lives > 10 years are averaging the radiation intensity over the 11 years cycle as well as over the period of the meteorite along its path. According to Kepler's second law the meteorites will spend most of their time near the aphelion, at solar distances of about 3 to 4 a.u. and thus the average along the path will depend primarily on the radiation level at those solar distances. Contrary to this, short lived isotopes ($T_{1/2} \leq 100$ days) are integrating the radiation only over a relatively short fraction of the meteorite orbit near the orbit of the earth. This feature enables us in principle to compare radiation intensity levels at around 1 a.u. with those at around 3 to 4 a.u., which will not be possible with artificial space probes for years to come.

A suitable pair of isotopes for this investigation are ^{37}A ($T_{1/2} = 35$ days) and ^{39}A ($T_{1/2} = 325$ years), because in iron meteorites their relative production rate depend very little on the energy. Thus depth effects do not constitute serious problems.

Unfortunately iron meteorite falls are very infrequent. Therefore, this method will have to be applied also to stones, although a quantitative evaluation here is more difficult, because of the complexity of the target material, mainly the presence of calcium.

STOENNER, SCHAEFFER and DAVIS [51] have measured in the stone meteorite (chondrite) Hamlet a ratio $^{37}\text{A}/^{39}\text{A} = 2.0 \pm 0.3$. From experiments with 3 GeV protons they obtain a production ratio of 1.2 ± 0.2 , within the limits of error this result is in agreement with a constant irradiation intensity along the orbit of this meteorite. Similar experiments of FIREMAN and DEFELICE [52] on the same stone and on the iron meteorite Aroos indicate somewhat higher $^{37}\text{A}/^{39}\text{A}$ ratios than could be expected on the basis of 2 GeV irradiation experiments.

In discussing these results two facts have to be born in mind. In the first place, the ^{37}A concentrations may be affected strongly by a single solar flare or a Forbush event, because of its short half-life. Therefore, the $^{37}\text{A}/^{39}\text{A}$ ratios would have to be measured in a considerable number of meteorites, before a comparison between the average fluxes in the region of the orbit of the earth and in the asteroidal belt becomes possible. Secondly, in stone meteorites, the $^{37}\text{A}/^{39}\text{A}$ ratio depends strongly on the calcium content, as calcium produces mainly ^{37}A (STOENNER, SCHAEFFER and DAVIS [51]). From this follows, that the $^{37}\text{A}/^{39}\text{A}$ production ratio in chondrites will depend appreciably on the energy, the production from calcium being relatively stronger at lower energies. Nevertheless, an important information is obtained from these experiments. There does not exist a very striking difference between the fluxes around 1 a.u. and around $(3 \div 4)$ a.u., if these meteorites have had orbits roughly similar to the orbit given in Fig. 2.

In the quiet solar period ahead events like solar flares and Forbush decreases will be much more seldom than they are now, and accurate comparative measurements of short- and long-lived isotopes in only a few meteorites should yield a reliable average. Combined with direct measurements on the radiation intensity in the vicinity of the orbit of the earth a figure for the radiation intensity at 3 to 4 a.u. solar distance will be made available.

From the tritium concentrations in small stone meteorites GEISS and OESCHGER [32] have calculated a mean cosmic ray flux of 0.6 to 0.8 particles/cm² s sr (with energies < 300 MeV) on the basis of measured tritium production cross-sections and estimates of shielding and secondary production effects. GOEBEL and SCHMIDLIN [53] with a similar method arrived at 0.75 particles/cm² s sr. These values represent averages over the solar cycle and over the period of the meteorite (cf. Section 7). They would indicate a total cosmic ray flux at $(2 \div 4)$ a.u., which is some 2 or 3 times higher than the flux observed at 1 a.u. One might look at these results as being in disagreement with the evidence from the $^{37}\text{A}/^{39}\text{A}$ ratio, and might tend to disregard them, because

they are not based on *ratios* of spallation product abundances. However, there actually does not exist a real disagreement. Both in stone and in iron meteorites, the low energy radiation at aphelion is very little recorded by ^{39}A in meteorites. In the Aroos iron the low energy is strongly absorbed and also has a low excitation function for ^{39}A . In the small stone Hamlet low-energy particles will produce ^{37}A from calcium, but very little ^{39}A . So it is quite possible that the observed high tritium concentration is indicative of a relatively strong low-energy component in the aphelion regions, which is not picked up by ^{39}A .

From the comparison of ^{45}Sc with ^{22}Ne abundances WÄNKE [48] also came to the conclusion, that the iron meteorites pick up more low-energy particles than there are in conventionally given spectra (cf. eq. (26)).

The difficulties involved in the estimate of total fluxes will be discussed in Section 10. The figures derived from both the tritium and the ^{45}Sc concentrations rely heavily on cross-section data and on estimates on the secondary production. In spite of this, the present evidence on meteorites seems to indicate that the flux at lower cosmic-ray energy in the aphelion region of meteorites is higher than it is estimated near the orbit of the earth.

9. - Long-time intensity variation ($(10^8 \div 10^9)$ years).

In the discussion of time variations of cosmic ray intensities perhaps of the greatest interest is the investigation of changes over times comparable to the life-time of individual galactic ray particles or even the age of the galaxy itself. In fact, the investigation of variations of the intensity over 10^9 years or more goes to the heart of the matter of theories on the origin of cosmic radiation.

The only spallation product in meteorites with a half-life sufficiently long to show radiation intensity variations beyond about 10^7 years is ^{40}K . In iron meteorites the chemical abundance of potassium is so low that cosmic ray-produced ^{40}K and ^{41}K can be detected (VOSHAGE and HINTENBERGER [54]; HONDA [55]).

As the ratio of the production cross-sections for ^{40}K and ^{41}K from iron may be expected to depend only very little on the energy, the ratio $^{40}\text{K}/^{41}\text{K}$ may be used for the detection of intensity variations without too much regard to the effects of shielding, *i.e.* the size of the meteorite etc.

For constant irradiation between the break-up time T_b and the time of fall (recent)—an assumption which includes the absence of relevant space erosion during this period—the measured concentration ratio will be (VOSHAGE and HINTENBERGER [54]):

$$(22) \quad \frac{^{40}\text{K}}{^{41}\text{K}} = \frac{\sigma_{^{40}\text{K}}}{\sigma_{^{41}\text{K}}} \frac{1 - \exp[-\lambda T_s]}{\lambda T_s}.$$

In this case the break-up age T_B is equal to the radiation age T_s as defined in eq. (2), which is deduced from a pair of a stable and a radioactive ($T_{\frac{1}{2}} < 10^7$ years) spallation isotope. In case of long-term variations of the radiation intensity or of appreciable space erosion one would observe a deviation from the relation between the ratio $^{40}\text{K}/^{41}\text{K}$ and T_s , as given above.

How sensitive a method do we have here for testing long-term intensity variations of irradiation? For a discussion of this question we have chosen an exponential time-dependence for the intensity:

$$(23) \quad I(t) = I_0 \exp[\gamma t].$$

If t counts backwards, $\gamma > 0$ is a decrease and $\gamma < 0$ an increase with (positive) time of the radiation intensity received at the location of the sample in the meteorite. We obtain

$$(24) \quad \frac{^{40}\text{K}}{^{41}\text{K}} = \frac{\sigma_{^{40}\text{K}}}{\sigma_{^{41}\text{K}}} \frac{\gamma}{\gamma - \lambda} \frac{\exp[(\gamma - \lambda)T_B] - 1}{\exp[\gamma T_B] - 1}.$$

In this general case the radiation age T_s as obtained from a stable and radioactive ($T < 10^7$ years) isotope is related to the break-up time T_B by (cf. eq. (3)):

$$(25) \quad T_s = \frac{1}{\gamma} (\exp[\gamma T_B] - 1).$$

In Fig. 11 theoretical $(^{40}\text{K}/^{41}\text{K})/(\sigma_{^{40}\text{K}}/\sigma_{^{41}\text{K}})$ values are plotted *vs.* radiation age T_s for the three different radiation histories: 1) $\gamma = 1/700$ million years; 2) $\gamma = -1/700$ million years and 3) constant irradiation. It should be noted that for changing irradiation intensity, T_s is not identical with the break-up time T_B . It may be seen from Fig. 11 that the $^{40}\text{K}/^{41}\text{K}$ method indeed is sufficiently sensitive to indicate intensity variations of a similar magnitude as those assumed here, if meteorites with radiation ages of 10^6 years and more can be included. A good check of the validity of this method will be to ensure that iron meteorites with young radiation ages give $^{40}\text{K}/^{41}\text{K}$ ratios which converge towards a constant universal value for $T_s \rightarrow 0$.

If $^{40}\text{K}/^{41}\text{K}$ values below the curve $I = I_0 = \text{constant}$ are found it will be difficult to distinguish between space erosion and a true intensity increase.

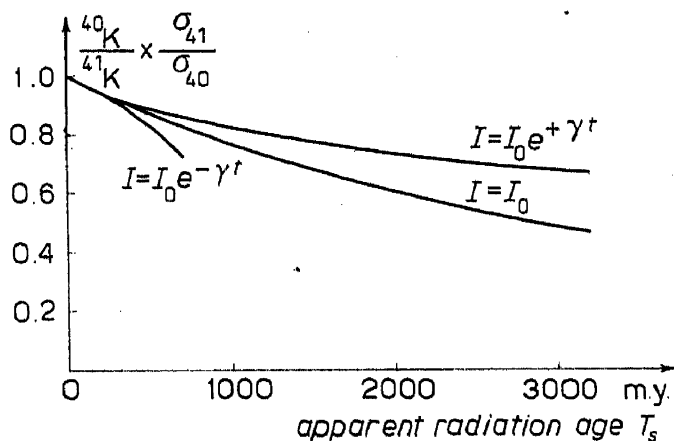


Fig. 11. — Expected $^{40}\text{K}/^{41}\text{K}$ ratios *vs.* radiation age for constant radiation intensity ($I = I_0$) and for variable intensities. $\gamma = 1/700$ million years.

Values above this curve are more unambiguously interpreted as intensity decreases, but any discussion of such results will have to consider changes which may have resulted from variations of the orbital elements of the meteorite.

VOSHAGE and HINTENBERGER [54] have been able to measure $^{40}\text{K}/^{41}\text{K}$ ratios in several iron meteorites, but there is not yet sufficient agreement on the radiation ages of these meteorites to allow a discussion of these data in terms of intensity variations of cosmic radiation at this moment.

10. – The spectrum and the absolute intensity of cosmic radiation.

Observed cosmic ray spectra are usually written in the following form

$$(26) \quad f(E) dE = \frac{f_0}{(1 + E)^\alpha},$$

where E is the kinetic energy in GeV and α is approximately 2.5 (cf. PETERS [56]; SINGER [57]; McDONALD and WEBBER [58]; ARNOLD, HONDA and LAL [47]). At low energies (< 1 GeV), however, the spectrum is not well known and it is changing with time (NEHER [59]; McDONALD and WEBBER [58]). Besides short-time variations, there exists a trend with the solar cycle. Low energy particles are more abundant in quiet than they are in the active periods. The other effect, the direct solar emission has opposite phase, and it is not yet really established if this solar output on average is able to compensate the other effect to any large extent. Space probe experiments have shown that these modulation effects have to be considered as affecting the whole inner part of the solar system and are not restricted to the neighbourhood of the earth (FAN, MEYER and SIMPSON [60]).

Whatever mechanism is suggested for the modulation of the galactic radiation (PARKER [61]; ELLIOT [62]) the effects are bound to decrease at some solar distance. So at 3 to 4 a. u., the aphelion region of meteorites, a radiation intensity may permanently exist which is as high or even higher than the intensity received at 1 a. u. during quiet periods. Meteorites, therefore, can be expected to monitor an average radiation spectrum which is much closer to the galactic spectrum than that received at the neighbourhood of the earth.

We shall discuss in the following, how different spectra are reflected in the spallation product distribution of an infinitely small iron meteorite. As before, we shall base this discussion on the tenuous assumption that we deal with irradiation by one type of particles, *i.e.* protons. Actually, the abundance of α -particles is such, that their effect cannot be neglected. However, their spectrum seems to be so similar to that of the protons, that the mixture of protons and α -particles can be represented by a proton spectrum.

For energies larger than about 3 to 5 GeV, there is little question that eq. (26) coincides with the cosmic ray spectrum at 1 a.u. It is reasonable to assume that this is true also at a solar distance of about 4 a.u. The real question concerns the fluxes of the true galactic radiation at lower energies, below 2 GeV. Therefore, in the discussion of spallation products in small meteorites we cannot limit ourselves to a spectrum of the form (26), but we have to keep considerable flexibility in the low energy range. A spectral form

$$(27) \quad f(E) dE = f_0 \frac{\exp [-\beta E^{-\frac{1}{2}}]}{E^\alpha} dE,$$

is very suitable for our purposes for two reasons. In the first place, the relative abundance of low energy particles can be determined through the choice of β (Fig. 12). If $\alpha = 2.5$ is taken, spectrum (27) approaches spectrum (26) at high energies. If $\beta = 1.7$ is chosen, the two spectra (27) and (26) agree within 20% for $E \geq 300$ MeV. Secondly, eq. (5) can be integrated directly if the energy spectrum is of the form (27). The choice of this general spectral type practically does not constitute any limitation, because with a proper choice of β observed spectra (cf. McDONALD and WEBBER [58]) often may be approximated even better by (27) than by (26).

The ΔA distribution obtained from irradiation with the spectrum (27) is given by integrating eq. (5):

$$(28) \quad N(\Delta A) = \frac{f_0 k_0 e^{-(n-1)} \Gamma(n)}{(\Delta A + \beta/e)^n},$$

with (cf. eqs. (10), (11) and (12))

$$(29) \quad n = \frac{1}{2}(3\alpha - 1).$$

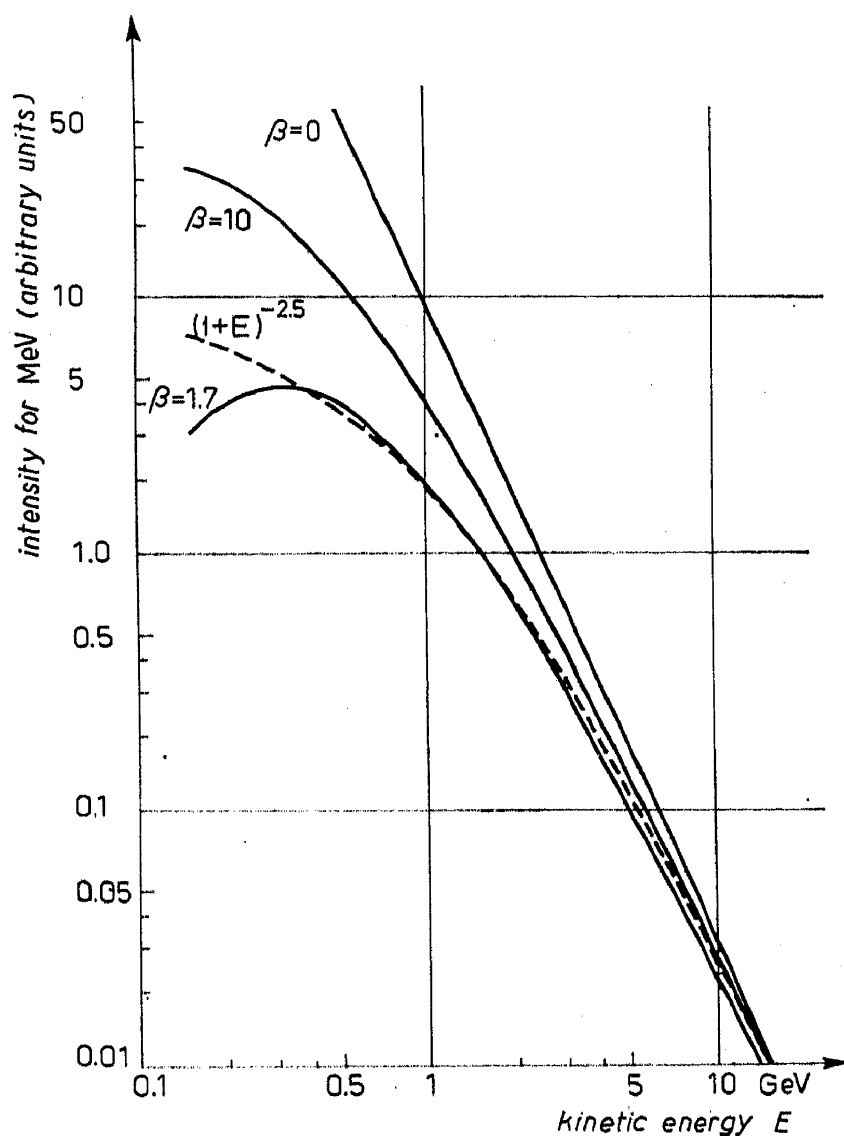


Fig. 12. - Possible cosmic ray energy spectra. Solid lines give spectra of the type $f(E)dE = f_0 \cdot E^{-\alpha} \cdot \exp [-\beta E^{-\frac{1}{2}}] dE$, with different β 's representing contributions of low energy particles. Dashed line gives spectrum $f(E)dE = f_0(1 + E)^{-2.5} dE$.

These spallation product distributions $N(\Delta A)$ are plotted in Fig. 13 for the different spectra given in Fig. 12. It is seen that the different assumed spectra do lead to quite different spallation product distributions. The analysis of stable spallation product abundances from ^{21}Ne to ^{50}V in a very small iron meteorite would afford a very sensitive test on the average radiation spectrum

received by the meteorite during its life-time in space. If radioactive spallation products can be measured with sufficient accuracy, the spectrum would be tested for shorter periods. If we consider only $6 \leq A \leq 35$, the spectrum is tested effectively only down to about 150 MeV.

In the iron meteorites Grant and Aroos with post-atmospheric sizes of 480 and 180 kg respectively we have determined the irradiation spectrum

$$(16) \quad f(E) dE = \frac{f_0}{E^{1.93}} dE.$$

This spectrum is effectively determined by the spallation products only in a limited energy range, which may be given roughly as $200 \text{ MeV} \leq E \leq 10 \text{ GeV}$.

In meteorites of this size the radiation spectrum is affected by ionization and nuclear interactions. Primary particles below about 500 MeV are absorbed by ionization. Thus the information on the low energy part of the primary radiation is lost. Even for high energies it is very difficult to work back from the effective spectrum to the primary spectrum, because not only the above mentioned effects have to be taken into account quantitatively but also pre-atmospheric size and shape have to be adopted.

Actually, spectrum (16) is not so different from the generally adopted primary energy spectrum (26). Between 800 MeV and 15 GeV the two spectra agree within about 20 percent.

From the ΔA distribution, a radiation age can be derived corresponding to the age defined in eq. (2) which is based on the weighted average concentrations of all the stable and all the radioactive spallation

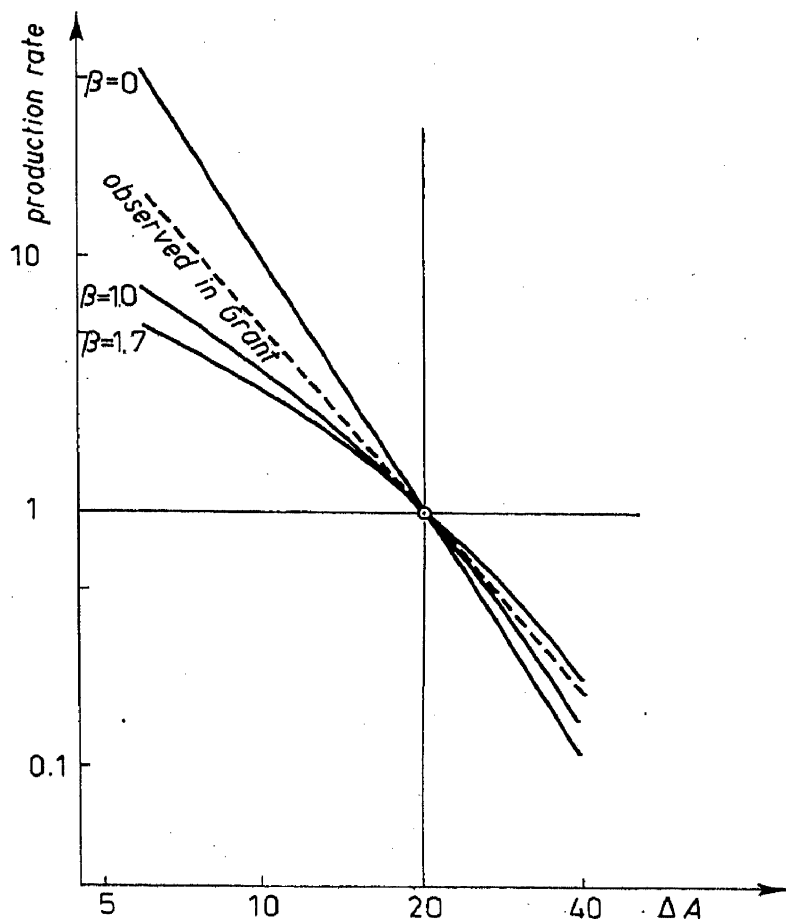


Fig. 13. — ΔA -distribution obtained from different spectra given in Fig. 12. The distribution observed in the Grant meteorite is included.

products. If

$$(30) \quad N_s(\Delta A) = \frac{a_s}{(\Delta A)^n}$$

and

$$(31) \quad N_i(\Delta A) = \frac{a_i}{(\Delta A)^n}$$

are the measured distributions of the stable and the radioactive spallation products respectively, the radiation age is given by

$$T_s = \frac{a_s}{a_i}.$$

If a_s is measured in 10^{-8} cm³/g and a_i in dpm/kg

$$T = 510 \frac{a_s}{a_i} \cdot 10^6 \text{ y.}$$

is obtained.

In principle, from the concentrations of radioactive spallation products the absolute intensity of cosmic radiation (the time average along the orbit of the meteorite) can be deduced. Contrary to all applications discussed so far, direct calculations of intensity figures cannot be based on *ratios* of spallation products, and effects of shielding or secondary production cannot be canceled out. These effects together with the cross-section data adopted enter directly into the calculated absolute intensity, so that relevant data can be obtained only from relatively small meteorites.

If the intensity of cosmic radiation in the region of the asteroidal belt is different from the intensity we have at one a.u., this most probably will be mainly due to a difference at lower energies. Therefore, in determining the total flux, one has to simultaneously determine the energy spectrum.

Because of the lack of figures from relatively small iron meteorites, we shall calculate here only the actual irradiation intensity at the site of the Aroos sample investigated by HONDA and ARNOLD [50].

From the irradiation spectrum $f(E) dE = f_0 E^{-\alpha} dE$ the integral spectrum

$$(33) \quad F(E) = \int_E^{\infty} f(E) dE = f_0 \frac{E^{-(\alpha-1)}}{\alpha-1}$$

is obtained. f_0 can be expressed in terms of a_i , using the equations

$$(34) \quad f_0 = \frac{a_i}{k_0 c^{-(n-1)} \Gamma(n)}.$$

Figure 8 yields $a_i = 2.39 \cdot 10^4$ dpm/kg for the Aroos meteorite sample, giving $f_0 = 1.05$ (E in GeV), and

$$(35) \quad F(E) = 1.13 E^{-0.93} \quad (E \text{ in GeV})$$

as proton equivalent irradiation flux at the site of the sample. Equation (35) can be used only for $200 \text{ MeV} \leq E \leq 6 \text{ GeV}$, because at lower energies the spectrum is not really determined by the spallation product distribution which is used. The uncertainty at high energies in the differential spectrum can be disregarded in the integration. The irradiation flux above 1 GeV at the site of the Aroos sample is $F(1 \text{ GeV})/4\pi = 0.09$ protons (equival.)/ $\text{cm}^2 \text{ s sr}$.

TABLE I. — Characteristic energies (in GeV) for the production of an isobar ΔA from irradiation spectrum $f(E) dE = f_0 E^{-1.97} dE$. E_m = energy of maximum production, E_p = median energy. 60 percent of the isobar is produced between E_m and $2E_p$.

ΔA	E_m	E_p	$2E_p$
10	0.17	0.51	1.02
20	0.52	1.53	3.06
30	0.95	2.75	5.50

REFERENCES

- [1] F. G. HOUTERMANS: *Suppl. Nuovo Cimento*, **11**, 390 (1954).
- [2] W. F. LIBBY: *Radiocarbon Dating*, 1st Ed. (Chicago 1952).
- [3] W. F. LIBBY: *Radiocarbon Dating*, 2nd Ed. (Chicago 1955).
- [4] G. L. KULP and M. L. VOLCHOK: *Phys. Rev.*, **90**, 713 (1953).
- [5] J. N. ROSHOLT, C. EMILIANI, J. GEISS, F. F. KOCZY and P. J. WANGERSKY: *Journ. Geol.*, **69**, 162 (1961).
- [6] H. DE VRIES: *Researches in Geochemistry*, H. ABELSON (Ed.) (New York, 1959).
- [7] H. E. WILLIS, H. TAUBER and K. O. MÜNNICH: *Amer. Journ. Sci., Radiocarbon Suppl.*, **2**, 1 (1960).
- [8] H. E. SUESS: Preprint (1961).
- [9] B. PETERS: *Proc. Indian Acad. Sci.*, **41**, 67 (1955).
- [10] J. R. ARNOLD: *Science*, **124**, 584 (1956).
- [11] F. BEGEMANN: *Zeits. f. Naturfor.*, **14 a**, 334 (1959).
- [12] S. EPSTEIN: *Courses on Nuclear Geology* (Pisa, 1961).
- [13] D. LAL and B. PETERS: Preprint (1961).
- [14] H. CRAIG: *Tellus*, **9**, 1 (1957).
- [15] N. W. RAKESTRAW, H. E. SUESS and H. OESCHGER: *Proc. I. Internat. Oceanographic Congress* (New York, September 1959).

- [16] G. S. BIEN, N. W. RAKESTRAW and H. E. SUESS: *Tellus*, **12**, 436 (1961).
- [17] W. S. BROECKER, R. GERARD, M. EWING and B. C. HEEZEN: *Journ. Geophys. Research*, **65**, 2903 (1960).
- [18] J. ZÄHRINGER: this volume, p. 279.
- [19] P. EBERHARDT and J. GEISS: *Proceedings* (September 1960).
- [20] H. C. UREY: *Progress in Physics and Chemistry of the Earth*, **2** (London, 1957).
- [21] E. ANDERS and G. G. GOLES: *Journ. Chem. Education*, **38**, 58 (1961).
- [22] Z. CEPLECHA: *Bull. Astron. Inst. Czechoslovakia*, **12**, 21 (1960).
- [23] C. A. BAUER: *Phys. Rev.*, **72**, 354 (1947).
- [24] J. H. HOFMAN and A. O. NIER: *Phys. Rev.*, **112**, 2112 (1958).
- [25] P. SIGNER and A. O. NIER: *Journ. Geophys. Research*, **65**, 2947 (1960).
- [26] J. GEISS, H. OESCHGER and P. SIGNER: *Zeits. f. Naturfor.*, **15 a**, 1016 (1960).
- [27] K. GOEBEL P. SCHMIDLIN and J. ZÄHRINGER: *Zeits. f. Naturfor.*, **14 a**, 996 (1959).
- [28] H. WÄNKE: Private communication (1960).
- [29] D. HEYMAN and O. A. SCHAEFER: *Amer. Geophys. Union Meeting* (Washington, April 1961).
- [30] E. L. FIREMAN and J. DE FELICE: *Geochim. et Cosmochim. Acta*, **18**, 183 (1960).
- [31] F. L. WHIPPLE and E. L. FIREMAN: *Nature*, **183**, 1315 (1959).
- [32] J. GEISS and H. OESCHGER: *Space Research*, H. K. KALLMAN-BIJL Ed., in *Proc. I. Internat. Space Sci. Symp.*, Nice, January 1960 (Amsterdam, 1960), p. 1071.
- [33] H. C. UREY: *Journ. Geophys. Research*, **64**, 1721 (1959).
- [34] F. BEGEMANN, J. GEISS and D. C. HESS: *Phys. Rev.*, **107**, 540 (1957).
- [35] P. EBERHARDT, J. GEISS and H. LUTZ: *Schweiz. Phys. Ges. Symp.* (Rolle, May, 1961), *Helv. Phys. Acta*, in press.
- [36] E. L. SPRENKEL: Brookhaven Nat. Lab., Thesis, unpublished (1959).
- [37] D. E. FISHER and O. A. SCHAEFFER: *Geochim. et Cosmochim. Acta*, **20**, 5 (1960).
- [38] H. C. VAN DE HULST: *Astrophys. Journ.*, **105**, 471 (1947).
- [39] H. ELSÄSSER: *Mittlg. Astron. Inst., Tübingen*, **35**, 61 (1958).
- [40] E. ANDERS: *Geochim. et Cosmochim. Acta*, **19**, 53 (1960).
- [41] J. A. O'KEEFE: *Space Research*, H. K. KALLMAN-BIJL (Ed.), *Proc. I. Internat. Space Sci. Symp.* (Nice, January 1960) (Amsterdam, 1960), p. 1080.
- [42] T. GOLD: *Mon. Not. Royal Astron. Soc.*, **115**, 585 (1955).
- [43] G. RUDSTAM: Univers. of Uppsala, Thesis (1956).
- [44] D. W. BARR: Univers. of Calif., Thesis, UCRL-3793 (1957).
- [45] M. HONDA and D. LAL: *Phys. Rev.*, **118**, 1618 (1960).
- [46] H. OESCHGER and U. SCHWARZ: unpublished (1961).
- [47] J. ARNOLD, M. HONDA and D. LAL: Preprint (1961).
- [48] H. WÄNKE: *Zeits. f. Naturfors.*, **15 a**, 953 (1960).
- [49] M. HONDA and H. STAUFFER: private communication (1961).
- [50] M. HONDA, SHEDLOVSKY and J. ARNOLD: *Geochim. et Cosmochim. Acta*, **22**, 133 (1960).
- [51] R. W. STOENNER, O. A. SCHAEFFER and R. DAVIES: *Journ. Geophys. Research*, **65**, 3025 (1960).
- [52] E. L. FIREMAN and J. DE FELICE: *Journ. Geophys. Research*, **65**, 3035 (1960).
- [53] K. GOEBEL and P. SCHMIDLIN: *Geochim et Cosmochim. Acta*, **17**, 342 (1959).
- [54] H. VOSHAGE and H. HINTENBERGER: *Zeits. f. Naturfors.*, **14 a**, 828 (1959).
- [55] M. HONDA: *Geochim et Cosmochim. Acta*, **17**, 148 (1959).
- [56] B. PETERS: *Progress Element. Particle and Cosmic Ray Phys.*, **1**, Ch. IX (1952).

- [57] S. F. SINGER: *Progress Element. Particle and Cosmic Ray Phys.*, **4**, Ch. IV (1958).
- [58] F. B. McDONALD and W. R. WEBBER: *Space Research*, H. K. KALLMAN-BIJL (Ed.), *Proc. I. Internat. Space Sci. Symp. Nice, January 1960* (Amsterdam, 1960), p. 968.
- [59] H. V. NEHER: *Phys. Rev.*, **103**, 228 (1956).
- [60] C. Y. FAN, P. MEYER and J. A. SIMPSON: *Phys. Rev. Lett.*, **5**, 269 and 272 (1960).
- [61] E. N. PARKER: *Phys. Rev.*, **110**, 1445 (1958).
- [62] H. ELLIOT: this volume (p. 16).

Some Recent Investigations on Primordial Rare Gases and ^{129}Xe in Meteorites.

J. ZÄHRINGER

Max-Planck-Institut für Kernphysik - Heidelberg

Introduction.

Meteorites have a different history from that of terrestrial matter. During their flight in space they are exposed to cosmic rays, solar particles and cosmic dust, and are continuously influenced by processes which cannot affect terrestrial material through its shielding atmosphere.

Cosmic rays produce spallation products in meteorites. This process is integrated over a long period of time, and from the isotopic variations one can deduce the history of cosmic rays. Also the duration of the exposure to cosmic rays or changes of the size of the meteoritic body itself by breaking up, or by space erosion, can be determined.

Meteorites differ also in another aspect from terrestrial matter. As seen from the isotopic variations of some decay products of natural radioactivities, meteorites represent the oldest material so far analysed. One can expect that also some traces of events in the early history of our planetary system are left. In fact during the last few years a number of isotopic variations have been found, which may reveal some information on the « prehistory » of meteorites and perhaps of our solar system.

It is the purpose of this article to treat these « prehistoric » isotopic variations and to present the current explanations and conclusions of this new field in meteorite research.

1. – Isotopic variations in meteorites.

Since Prof. GEISS reports mainly on cosmic ray effects during the last 10^9 years, this subject will not be discussed here in detail and no references will be given. Only a short introductory survey is given of all isotopic variations which cannot be understood by chemical fractionation processes.

Generally the isotopic composition of the elements in meteorites, as analyzed so far, is the same as in terrestrial elements. One finds large variations, however, in those elements, which are only present in traces below concentrations of about 10^{-8} g/g. These changes are due to processes which add isotopes in different composition and can therefore affect isotopic ratios considerably.

Several elements with such isotopic enrichment have been found in meteorites during the last 10 years. The processes responsible for these changes are summarized, together with the main results, as follows:

1'1. *Natural radioactivity.* – Natural radioactive elements such as ^{40}K , ^{87}Rb , ^{235}U , ^{238}U , ^{232}Th and ^{187}Re have a half-life of the order of 10^9 to 10^{10} years and the daughter products cause large enrichments in the corresponding isotopes ^{40}Ar , ^{87}Sr , ^{207}Pb , ^{206}Pb , ^{208}Pb , ^4He and ^{187}Os . The isotopic variations are mostly higher than in terrestrial material, due to the greater age of meteorites. From these isotopic changes an age of $4.5 \cdot 10^9$ years has been derived for meteorites as well as for the age of our solar system. Unfortunately all methods imply assumptions, which are based on a simple origin of all meteorites. Therefore these assumptions as well as the experimental results need further investigation before a final time scale for meteor forming processes can be given. There is good experimental evidence that the age of $4.5 \cdot 10^9$ years is too low.

1'2. *Spallation products.* – Out in space, meteorites are exposed to the primary cosmic radiation and to solar particles. The high-energy particles produce reaction products in the meteorites, such as evaporation particles, spallation, fragmentation and fission products. For a given target nucleus

TABLE I. – *Spallation-produced rare gas isotopes in iron*

	$^4\text{He} \cdot 10^{-8} \text{ cm}^3/\text{g}$	$^3\text{He} \cdot 10^{-8} \text{ cm}^3/\text{g}$	$^3\text{He}/^4\text{He}$	$^{21}\text{Ne} \cdot 10^{-8} \text{ cm}^3/\text{g}$
Williamstown	1920	480	0.25	6.0
Arispe	930	270	0.29	2.3
Odessa	710	220	0.31	2.3
Sikhote-Alin	481	130	0.27	1.5
Toluca	295	100	0.34	0.8
Canon Diablo	42	4.2	0.10	0.11?

all elements with lower atomic weight are formed. The yield curve depends on the chemical composition, on the energy of the incident particle and on the depth of the sample within the meteoritic (pre-atmospheric) body. The production of evaporated particles with $A < 5$ is not very different for iron and stone meteorites. The distribution of residual nuclei, however, is dependent mainly on the ratio of iron to silicate phase. The isotopic abundances of spallation-produced elements is roughly given by the stability of isotopes and can be estimated from the parabolic energy valley. Thus spallation-produced isotopes do not have very large differences in abundance and can generally be distinguished from the natural isotopic composition. Table I shows the spallation-produced rare gas isotopes, measured in some iron meteorites [1].

The total amount of spallation products as found in meteorites is small, of the order of $(10^{10} \div 10^{13})$ atoms per g. Therefore only elements of very small concentration such as the rare gases, Sc and K in iron meteorites, and radioactive products have been detected. A list of all these isotopes is given in Table II.

From the amount of a radioactive isotope one can get some information on the cosmic ray intensity, since in the radiation equilibrium the production rate A is given by

$$A = N \cdot F \cdot \sigma.$$

(A decays/s, N = number of target atoms, F = flux of particles/s cm^2 , σ = production cross-section in cm^2). If σ is known, A is a measure of the flux about 3 half-lives ago, and with a suitable set of radio-nuclides as ^{22}Na , T , ^{36}Cl and ^{40}K the mean intensity change during the past can be detected. For this no

meteorites (cor = corrected for atmospheric contamination).

$^{20}\text{Ne}; ^{21}\text{Ne}; ^{22}\text{Ne}$	$^{36}\text{A} \cdot 10^{-8}$	$^{38}\text{A} \cdot 10^{-8}$	$^{40}\text{A} \cdot 10^{-8} \text{ cm}^3$	$^{36}\text{A cor}$	$^{38}\text{A cor}; ^{36}\text{A cor}$
1:0.8:0.9	17.6	26.4	73.6	17.4	1.5
1:1 :1	5.5	8.9	13.7	5.5	1.6
1:0.9:1	7.0	12.0	10.1	6.7	1.6
—	2.8	4.1	?	2.8	1.5
—	1.7	2.4	15	1.7	1.5
—	0.27	0.44	16.7	0.27	1.7

TABLE II. — *Stable and radioactive spallation-produced isotopes, which are so far detected in meteorites.*

	$^3\text{H} \rightarrow$	^3He	^4He	^{10}Be	^{20}Ne	^{21}Ne	^{22}Ne	$\leftarrow ^{22}\text{Na}$	^{26}Al	^{32}Si	$^{36}\text{Cl} \rightarrow$	^{36}Ar	^{37}Ar	^{38}Ar	^{39}Ar
<i>T</i>	12 y	—	—	$2.5 \cdot 10^6$ y	—	—	—	2.58 y	$7.4 \cdot 10^5$ y	700 y	$3.1 \cdot 10^5$ y	—	35 d	—	325 y
	^{40}K	^{41}K	^{44}Ti	^{45}Ca	^{45}Sc	^{46}Sc	^{48}V	^{49}V	^{51}Cr	^{53}Mn	^{54}Mn	^{56}Co	^{57}Co	^{60}Co	^{59}Ni
<i>T</i>	$1.3 \cdot 10^9$ y	—	200 y	164 d	—	84 d	16 d	330 d	28 d	$2 \cdot 10^6$ y	291 d	77 d	270 d	5.2 y	$8 \cdot 10^4$ y

absolute numbers would be needed since the yield curve in dependence of the atomic weight could be measured with the stable spallation products. If, however, the depth of the sample, and therefore the energy and number of secondaries of the particles change during time, the evaluation of these measurements becomes very difficult. The experimental results up to now indicate that the cosmic ray intensity has not changed more than a factor of two during the last 10^9 years. From the T-activity one gets a cosmic ray flux of about 0.6 particles/cm² s sr.

Of special interest are pairs of a radioactive and a stable daughter product of the rare gases, such as T- ^3He , ^{22}Na - ^{22}Ne , ^{36}Cl - ^{36}Ar and ^{40}K - ^{40}Ar . From the ratios $^3\text{He}/\text{T}$ etc., one can derive the time of exposure to cosmic rays. This age represents the true duration of irradiation, if again the depth of the considered sample has remained constant. If there is continuous breaking up or space erosion, these radiation ages represent nothing but a certain mechanical stability. In fact, the radiation ages of stony meteorites do vary within at least a factor of ten and are about 10^7 years whereas the radiation age of iron meteorites always lies around some 10^8 years. This indicates a systematic difference between these two kinds of meteorites. Besides the difficulties in explaining the present results, one experimental difficulty also has to be mentioned. Two iron meteorites of recent falls have been investigated by several authors for T and other spallation products. Normal yields have been found for all isotopes except for T, which could not even be detected. This difficulty also remains to be explained.

1'3. *Primordial rare gases.* — In 1956 GERLING and LEVSKII [2] detected in the stony meteorite Staroe Pesjanoe a very large amount of He, Ne and Ar, which they could not explain in any other way than by assuming that these gases had been trapped from a primordial source, while the meteorite was formed. This observation has been considered as somewhat exceptional. In 1960 ZÄHRINGER and GENTNER [3] found a very similar result in the achondrite Kapoeta. The ^4He content of both achondrites is higher by a factor of 1000. The isotopic composition of the neon and argon is similar to the

atmospheric neon and argon, however different enough to denote a different origin. In addition their abundance (number of atoms/number of silicon atoms) comes closer to what is called cosmic abundance, which is derived from star analysis or nuclear systematics.

Another distinct group of stony meteorites, having excess of rare gases not explainable by the spallation process or natural radioactivity are the carbonaceous chondrites, the enstatite chondrites and several ordinary chondrites. These meteorites contain mainly the heavier rare gases A, Kr and Xe and have in contrast to the two achondrites no He and Ne excess.

One is tempted to call these gases primordial, in order to indicate that these meteorites might be samples of planetary matter from which the solar system originally was formed.

1.4. *Extinct radioactivity.* — Another important isotopic variation has been detected by REYNOLDS [4] in the Richardton chondrite. He found an appreciable ^{129}Xe excess of about 50% and the most likely source of ^{129}Xe is ^{129}I , which decays with $T=17\cdot 10^6$ years in ^{129}Xe . Here we have the decay product of a very short-lived radioactive isotope, which is extinct or not detected so far and this ^{129}Xe excess has to be considered as a very important key to the early history of our solar matter.

If the ^{129}I decayed in the present meteorite sample, one can calculate, from the $^{129}\text{Xe}/^{129}\text{I}$ ratio, the time elapsed since the end of nucleosynthesis and the formation of the meteorite. REYNOLDS thus found for this time interval about $100\cdot 10^6$ years.

If the ^{129}I has decayed elsewhere this isotopic effect indicates that condensation of matter occurred before a complete mixing of all elements. In both cases the ^{129}Xe excess is a fascinating discovery and has therefore attracted wide scientific interest.

The search for products of other extinct radionuclides is under fast progress. The ^{205}Tl is the decay product of ^{205}Pb with the half-life of $35\cdot 10^6$ years. Meteoritic Tl however was not found to be anomalous [5], ^{107}Ag is a daughter product of ^{107}Pd , which decays with a half-life of $4.5\cdot 10^6$ years. Here MURTHY [6] has found in Ag extracted from iron meteorites an excess of about $(2\div 3)\%$ in the $^{107}\text{Ag}/^{109}\text{Ag}$ ratio.

These decay products of solid elements are also of great importance, since they are not so sensitive to heating losses as gases.

The decay products of extinct radioactivities can be considered as a sensitive clock of the very early events of the meteorite forming processes. It reads millions of years at the end of a timescale of several 10^9 years and may date the prehistory of meteorites or perhaps also of planetary matter.

The decay of extinct radionuclides furnishes also a potential heat source, which may have been an importance factor in the evolution of meteorites and

planetary matter [7, 8]. A new hypothesis for the creation of meteorites in small asteroidal bodies with extinct radioactive heating has been presented by FISH, GOLES and ANDERS [9].

In this article the present situation about primordial gases and extinct radioactivity is described. In addition results of some of the Author's experiments on this subject are included. The combination of these two kinds of isotope effects has been chosen since there seems to be a relation between ^{129}Xe excess and primordial gases. ^{129}Xe has so far always been observed in company with other Xe isotopes.

2. - Experimental results of primordial gases.

2.1. *Achondrites*. - In the previous chapter the rare gas isotopes of typical spallation products in iron meteorites have been shown, where a contribution of the processes 3) and 4) cannot be observed.

In Table III the rare gas content of a typical stony meteorite Norton County is presented. It contains a very high amount of spallation products but also some ^4He and ^{40}Ar from the U, Th and ^{40}K decay.

TABLE III. - The primordial rare gas content and the K-A-ages of Staroe Pesjanoe and Kapoeta and for comparison the ordinary rare gas content of Norton County (in $10^{-8} \text{ cm}^3/\text{g}$).

Meteorite	^3He	^4He	^{20}Ne	^{21}Ne	^{22}Ne	^{36}Ar	^{38}Ar	^{40}Ar	K in ppm	K-A-ages in 10^9 y
Norton County (Howardite)	215	1560	54	58.2	63.3	1.63	2.28	690 ± 50	65 ± 5	5.1 ± 0.25
Staroe Pesjanoe (Howardite)	252 (70)	705 000	2180 (23)	30.5 (25)	206 (27)	112 (~ 1)	21.3 (1.5)	2700 ± 100	363 ± 20	4.5 ± 0.2
Kapoeta	57,5 (7)	223 000	2340 (1.4)	7.8 (1.5)	180 (1.6)	88	17.2	1140 ± 50	165 ± 10	4.4 ± 0.15
Cosmic abundance Si = 10^3	?	3 000 000	7740	26	836	126	24	?	—	—

In the same table the rare gas content of the achondrites Staroe Pesjanoe and Kapoeta are listed and the difference in the ^4He and ^{20}Ne content is quite striking. From the low ^{21}Ne content in Kapoeta one can also see that the contribution of spallation must be small. Kapoeta thus represents a stony meteorite with nearly pure primordial rare gas content with the exception of some

radiogenic ^{40}Ar . Assuming a similar ratio, $^{22}\text{Ne}/^{21}\text{Ne} = 34$ in primordial and atmospheric neon only $1.5 \cdot 10^{-8} \text{ cm}^3/\text{g}$ ^{21}Ne is left for spallation and this also means that 90% of the ^3He content in Kapoeta is primordial. A similar splitting can be applied on Staroe Pesjanoe and the estimated spallation contribution is shown in brackets.

Under these conditions the $^4\text{He}/^3\text{He}$ ratio of primordial helium is in both meteorites about 4000. This is the first direct measurement of ^3He of a primeval source.

In both samples, after the spallation correction the $^{20}\text{Ne}/^{22}\text{Ne}$ ratio is about 13.5 and 35% higher than in the atmospheric neon. However no large difference exists in the $^{36}\text{Ar}/^{38}\text{Ar}$ ratio for both meteorites compared to 5.35 in the atmosphere [13, 3].

The presence of ^3He in primordial gases is expected and can be understood by nuclear reactions. In the d-burning processes the reactions $d(p, \gamma)^3\text{He}$, $d(d, n)^3\text{He}$ or $d(d, p)\text{T} \rightarrow ^3\text{He}$ may produce ^3He . In the presence of high-energy protons ($> 20 \text{ MeV}$) $^4\text{He}(p, 2p)\text{T} \rightarrow ^3\text{He}$ or $^4\text{He}(p, pn)^3\text{He}$ reactions give ^3He . From star observations only upper limits for the $^4\text{He}/^3\text{He}$ ratio are known.

GREENSTEIN [10] found for the sun > 50 and WILSON [11] > 10 for M stars. The $^4\text{He}/^3\text{He}$ ratio in cosmic rays is about 2 and seems to be very low in this comparison.

The higher $^{20}\text{Ne}/^{22}\text{Ne}$ ratio could be understood if the earth has lost an appreciable amount of its earlier neon content by diffusion processes [12].

TABLE IV. — Rare gas content of separated minerals from Kapoeta in comparison to pulverized average samples in $10^{-8} \text{ cm}^3/\text{g}$.

No.	Minerals	Weight mg	^3He	^4He	^{20}Ne	^{21}Ne	^{22}Ne	^{36}Ar	^{38}Ar	^{40}Ar	K in ppm	K-A-age in 10^9 y
27	Average samples	105.8	57.2	223000	2340	7.8	180	88	—	1140	165 ± 10	4.4 ± 0.15
28		133.8	55.2	205000	2280	7.5	159	87	—	1070	157 ± 10	4.4 ± 0.15
131a	Olivine	116	≤ 9	3160	≤ 26	1.7	—	< 0.3	0.1	71	—	—
126	Olivine	93.6	≤ 13	1740	≤ 24	~ 2.4	—	< 1.0	0.2	145	—	—
131b	Pyroxene	14.9	≤ 7	1550	≤ 20	1.3	—	< 1.4	0.6	360	—	—
178	Plagioclase	27.1	≤ 5	1000	≤ 30	1	—	~ 2	~ 1	2400	256 ± 10	4.9 ± 0.2
168a	Plagioclase	12.6	≤ 12	2150	≤ 20	~ 2.6	—	< 2.1	—	3300	—	—
178	Glass { from 80 μm fraction	9.7	73	257000	2630	9.7	196	117	23	1800	320 ± 40	4.0 ± 0.4
168b		15.4	—	232000	2060	7.1	160	—	—	—	—	—
166	80 μm fraction	20.1	59	188000	1860	6.6	140	76	14.9	1010	—	—

^{20}Ne would diffuse more easily than ^{22}Ne . The same effect should also be observed on argon. It seems very likely, that the heavier rare gases are not subjected to such a diffusion mechanism, because they may have been trapped at low temperature during the separation of the volatile elements.

In order to understand how these gases have been trapped within the meteorite, a mineral separation has been tried, and with a Kapoeta sample four components analysed. Table IV shows the results. The minerals olivin, pyroxen and plagioklas do not

contain primordial gases, while the glass has an enriched concentration relative to the sample with 80 μm grain size, from which the glass was separated.

In addition a degassing experiment has been performed with a powdered sample ($< 60 \mu\text{m}$) of Kapoeta. The sample was heated at different temperatures for a convenient time and the extracted gas analysed. Figure 1 gives the results in $10^{-8} \text{ cm}^3/\text{g}$ per hour in dependance of $1/T$. This picture shows, that for all isotopes the primordial component diffuses out at relative low temperature and that the spallation products are released at higher temperatures. This can be seen from the decreasing $^4\text{He}/^3\text{He}$ and $^{20}\text{Ne}/^{21}\text{Ne}$ ratios.

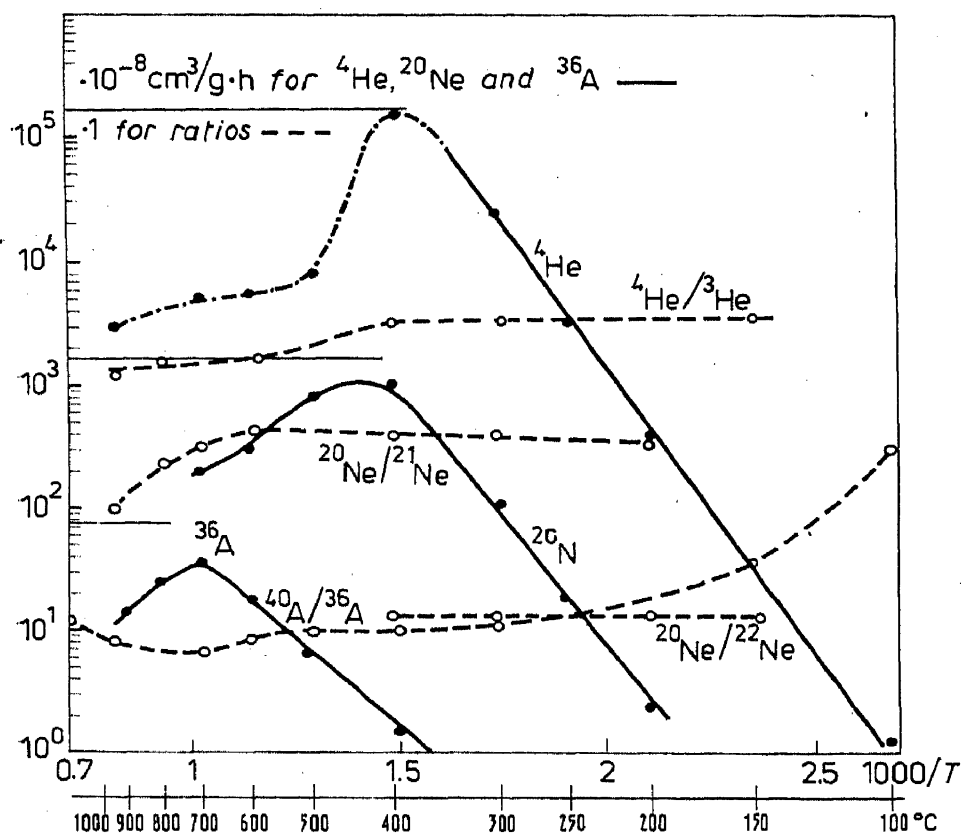


Fig. 1. — Rare gas release at different temperatures in $10^{-8} \text{ cm}^3/\text{g h}$ from the Kapoeta meteorite.

From this degassing experiment we also learned, that the temperature always must have been lower than 400 $^{\circ}\text{C}$, since it would be degassed in one hour at this temperature. Considering the age of $4.4 \cdot 10^9$ years the mean temperature has to be assumed to be below 0 $^{\circ}\text{C}$.

The degassing experiment of the Staroe Pesjanoe meteorite gave a similar result. The primordial gases were also released at lower temperatures and indicate a different trapping mechanism from that for spallation products.

The question of how these gases happened to come into this meteorite sample is not easy to answer. GERLING and LEVSKII [2] assumed that a high partial pressure of f.i. 2200 mm Hg of ^4He had been present when the me-

teorite solidificated. In the case of Kapoeta, the glassy particles have only trapped the rare gases and were probably later on baked together with the above mentioned minerals.

These glass particles could also have trapped the gas in a rare gas atmosphere as droplets, if they cooled very quickly. Since the size of the particles is very small ($< 100 \mu\text{m}$) this does not seem to be impossible. It is, however, difficult to imagine such a large mass to hold this atmosphere. Therefore one rather prefers that the meteoritic « parent bodies » incorporated these amounts of primordial gases and could not be degassed so completely during the melting and cooling of these bodies. Since the surface has been at low temperature, a tight silicate layer or an icy permafrost layer, as suggested by ANDERS, may have prevented the gases from escape. In the case of Kapoeta a volcanic origin seems very likely, where ejected glass solidified at the surface and was mixed with the crystalline components.

In this chapter the primordial rare gases of two achondrites have been discussed, whose amount as well as the isotopic composition is different from the terrestrial rare gases. There exists another group of meteorites, which also contain primordial rare gases, however, of a different composition. These meteorites are all chondrites and they will be treated in the following.

2'2. *Chondrites*. — Conclusive evidence that also chondrites may contain primordial gases has been given quite recently [3, 4, 13]. They have been found in the carbonaceous and enstatite chondrites in large concentration and to a smaller extent also in a series of ordinary chondrites. Already GEISS and HESS [14] noticed a somewhat higher ^{36}Ar content in two chondrites and supposed an additional unknown source.

In Table V the rare gas content of carbonaceous, enstatite and ordinary chondrites is listed. There seems to be a small amount of primordial He

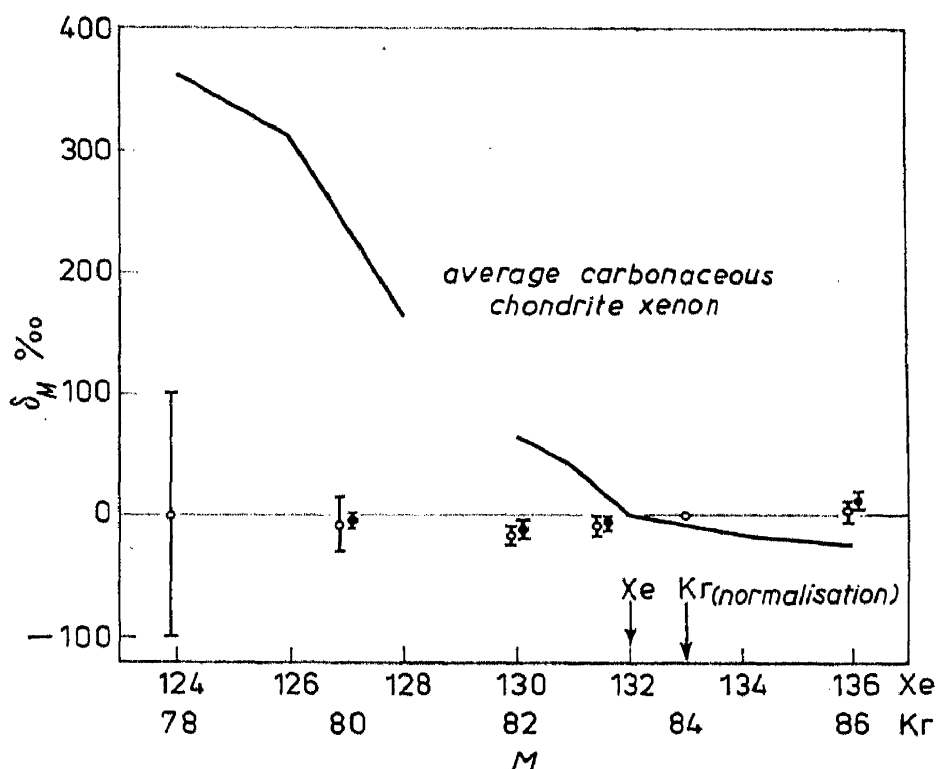


Fig. 2. — Krypton δ -values for carbonaceous chondrites displayed on the same plot with average xenon δ -values for carbonaceous chondrites [15]: ○ Murray krypton; ● Orgueil krypton.

TABLE V. — *Rare gas content of some carbonaceous,*

Meteorite		³ He	⁴ He	²⁰ Ne	²¹ Ne	²² Ne	³⁶ A	³⁸ A	⁴⁰ A
Mighei	Carb. Ch.	2.3	3 960	14.1	0.7	2.2	65.4	12.2	760
Bokkeveld	Carb. Ch.	0.5	2 450	10.7	0.09	1.5	46	9	(300)
Murray	Carb. Ch.	6.8	14 200	65.2	2.2	9.1	76.5	14.2	< 653
Abee	Enst. Ch.	12	1 320	10	2.46	3.3	37.2	6.9	6 900
St. Marks	Enst. Ch.	1.0	685	1.1	0.3	0.36	34.6	6.30	3 660
Indarch	Enst. Ch.	12.5	1 310	8.8	3.7	4.4	6.6	1.6	5 625
Forest Vale	Ch.	110	2 370	24	27.4	29.1	7.7	4.4	6 450
Bjurböle	Hyp. Ch.	19	1 550	4.75	4.85	5.15	3.9	1.2	6 010
Richardton	Br. Ch.	31.7	1 750	9.3	9.5	10.0	2.5	1.32	5 550
Nadiabondi		105	2 200	20	21	23	3.5	3.0	6 075
Miller	Br. O. Ch.	5.8	1 645	1.6	1.1	1.28	1.53	0.46	6 400
Forksville	Hyp. Ch.	48	1 300	9.2	9.65	10.3	1.75	1.21	5 700
Bruderheim	Br. O. Ch.	47	490	9.1	10.1	11.2	1.78	1.38	1 110
St. Isabel	Hyp. Ch.	64	1 790	11.9	11.9	12.8	1.84	1.75	6 000

and Ne only in some carbonaceous and enstatite chondrites. In general, however, only A and relatively high Kr and Xe values have been found in chondrites. The ³⁶A/³⁸A-ratio of carbonaceous chondrites and for enstatite chondrites this ratio is around 5.3 and not significantly different from atmospheric argon.

Also the krypton from the chondrites analysed does not show measurable deviation from terrestrial krypton (see Fig. 2). Here more measurements need to be made. Also the achondritic Kr and Xe has to be analysed in order to compare these two kinds of primordial gases.

An appreciable isotopic anomaly, however, has been found for xenon in these chondrites. All xenon isotopes from ¹²⁸Xe to ¹³²Xe are higher in the chondrites up to 40% and this is called the general anomaly [4].

statite and ordinary chondrites in 10^{-8} cm³/g.

⁸⁴ Kr	¹³² Xe	¹²⁹ Xe	¹²⁹ Xe/ ¹³² Xe	K ppm	K-A-age 10 ⁹ y	⁴ He-age 10 ⁹ y
—	—	—	—	450	< 2.2	—
—	—	—	—	445	< 1.2	—
0.7	1.1	1.15	1.05	380	< 2.8	—
0.15	0.094	0.60	6.4 (5.3)	822	4.7	3.7
0.14	0.048	0.240	4.4 (3.3)	1 275	2.97	2.1
0.036	0.064	0.25	4.0 (3.6)	910	4.2	2.9
0.06	0.072	0.11	1.48	810	4.62	4.5
0.016	0.046	0.84	1.82	840	4.43	3.8
0.021	0.024	0.041	1.70 (1.5)	812	4.36	4.0
0.02	0.020	0.029	1.44	790	4.56	4.2
0.013	0.014	0.021	1.5	720	4.8	4.1
—	0.009	0.016	1.85	765	4.50	3.4
< 0.013	0.013	0.017	1.30 (1.17)	890	1.85	—
< 0.02	0.008	0.016	2.0	817	4.48	3.8

In Fig. 2 the δ -values for Kr and Xe are plotted, where

$$\delta = 1000 \left[\left(\frac{\text{Xe}^m}{^{132}\text{Xe}} \right)_{\text{met.}} / \left(\frac{\text{Xe}^m}{^{132}\text{Xe}} \right)_{\text{terr}} - 1 \right],$$

is the deviation in ‰ of meteoritic from terrestrial xenon normalized at mass 132. The δ 's of ¹³⁴Xe and ¹³⁶Xe have somewhat negative values. In principle the Xe of all chondrites has this general anomaly and deviations can be explained by atmospheric contamination. This general anomaly does not inhibit the ¹²⁹Xe excess which is much larger. This is called the special anomaly and will be discussed below.

In the following the present explanations for the general Xe-anomaly are reviewed:

a) KURODA [16] suggested first that this anomaly is due to a change of terrestrial xenon abundance and that the meteoritic xenon is normal. He normalized at mass 128 and the heavier terrestrial xenon isotope may have been increased by products from spontaneous fission of ^{244}Pu . This effect leads to a terrestrial excess of the isotopes from ^{132}Xe to ^{136}Xe . The anomaly of the lighter Xe-isotopes was later explained by CAMERON [17] by adding solar xenon to the earth atmosphere. The solar xenon has been exposed to a neutron irradiation during the condensing sun and might account for these anomalies. This possibility has been reinvestigated by KRUMMENACHER *et al.* [18] who found that unreasonably high yields for spontaneous ^{244}Pu fission would be needed for mass 131 and 132.

b) KRUMMENACHER *et al.* [18] rather believe that the general anomaly is due to a fractionation effect as discussed for neon and that the heavier

isotopes ^{131}Xe to ^{136}Xe have been enriched in the meteorites by spontaneous fission. In this explanation both meteoritic and terrestrial xenon would be anomalous. Figure 3 shows a logarithmic plot of the ratio of terrestrial to meteoritic xenon *vs.* mass. The logarithmic scale describes the dependence of the mass in diffusion processes. If only fractionation took place all the points should lie on the straight line. Thus the heavier isotopes must be enriched in the meteoritic xenon.

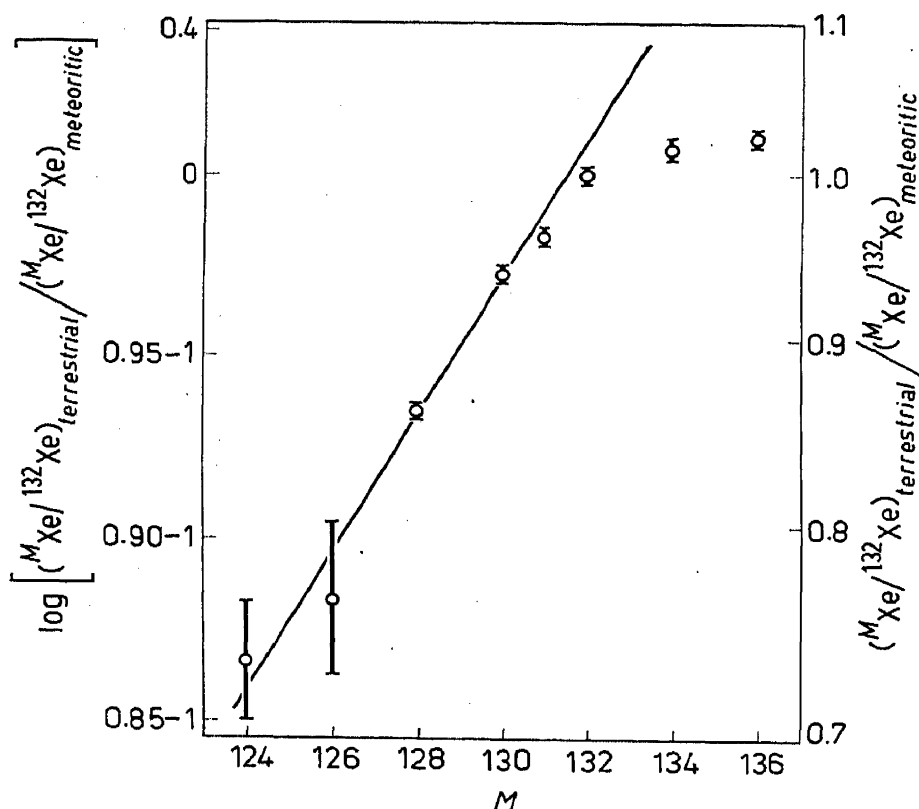


Fig. 3. - Logarithmic plot of terrestrial/meteoritic xenon *vs.* mass [15].

c) A further treatment of this problem has been published by FOWLER, GREENSTEIN and HOYLE [19]. These authors have been led to their hypothesis by the work of BONSAK and GREENSTEIN [20] who observed that the magnetic activity varies from star to star and that the amount of Li, Be, B produced varies similarly. They came to the conclusion that the D, Li, Be and B synthesis almost certainly occurred through spallation by high energy particles. The acceleration took place at the surface of

the condensing sun. The abundance of these elements indicates that the synthesis occurred during an intermediate stage of the solar system, when the planetary material was largely separated from hydrogen. The irradiation took place in planitesimals of the dimensions $(1 \div 50)$ m, consisting of meteoritic material embedded in an icy matrix. The surface layer of about 40 cm only has been irradiated. Secondary neutrons are thermalized in the icy matrix, and ${}^6\text{Li}$ and ${}^{10}\text{B}$ are depleted by (n, α) reactions.

In these processes also ${}^{124}\text{Xe}$, ${}^{126}\text{Xe}$, ${}^{128}\text{Xe}$ and ${}^{129}\text{Xe}$ should be produced by spallation, ${}^{134}\text{Xe}$ and ${}^{136}\text{Xe}$ enlarged by n-induced fission and ${}^{130}\text{Xe}$ and ${}^{132}\text{Xe}$ depleted by n-capture. In a detailed calculation these authors can explain the general anomaly of the Xe isotopes in meteorites.

At the moment it cannot be decided which model comes closest to the truth. The fact that krypton is of normal composition, is a difficulty in all these explanations.

2'3. ${}^{129}\text{Xe}$ excess or special anomaly. – The largest isotope variation in the chondritic xenon is caused by the ${}^{129}\text{Xe}$ excess, the so-called special anomaly. The ${}^{129}\text{Xe}/{}^{132}\text{Xe}$ ratio is 0.98 in atmospheric xenon. In the enstatite chondrites it is up to 6.5 and in ordinary chondrites between 1.2 and 2. In carbonaceous chondrites, where the primordial Xe content is the highest, the ${}^{129}\text{Xe}/{}^{132}\text{Xe}$ ratio is only slightly different from 1.

As for the source of this ${}^{129}\text{Xe}$ excess everyone is convinced that it is the decay product of ${}^{129}\text{I}$. ${}^{129}\text{I}$ decays, however, within $17 \cdot 10^6$ years and the question arises: *where was the ${}^{129}\text{I}$ produced?*

Three possibilities have been discussed so far:

1) Sudden nucleosynthesis.

2) Continuous nucleosynthesis.

3) High energy particle irradiation during an intermediate stage of the sun as described by FOWLER, GREENSTEIN and HOYLE [19] where ${}^{129}\text{Xe}$ is produced by spallation and by n-capture in ${}^{128}\text{Te}$.

The first possibility seems to be quite unreasonable under the aspects of modern nuclear synthesis. For the production in a continuous model the decay time seems to be unreasonably short for mixing the elements in the galaxies, transporting them to the solar system, forming the planetary disk and the meteorites. These difficulties would be avoided in the third model, where the production takes place *after* the material has been brought in our solar system.

Before applying the ${}^{129}\text{Xe}$ excess for any age determination model, a further crucial question has to be answered: *where did ${}^{129}\text{I}$ decay?*

Two of several possibilities may be mentioned:

a) ^{129}I decayed in the present meteorite sample and all excessive ^{129}Xe is due to ^{129}I . In this case the ^{129}Xe content should be proportional to the I content in minerals of differing I concentration.

b) ^{129}I decayed elsewhere in the so called protoplanetary cloud or in planitesimals as described above and the decay product ^{129}Xe was separated from iodine during chemical differentiation. Since different $^{129}\text{Xe}/^{132}\text{Xe}$ ratios are observed, the radiogenic ^{129}Xe could not have been completely mixed with primordial xenon because condensation might already have been in progress.

Any combination of these two alternative hypotheses may also be valid and cannot be excluded in the present state of knowledge.

REYNOLDS [4] assumes that the ^{129}I decayed in the present meteorite sample and the ^{129}Xe is considered as its decay product. The storage of the ^{129}Xe started when the meteorite was formed. With some assumptions about the $^{129}\text{I}/^{127}\text{I}$ ratio at the end of the iodine forming processes, one is able to deduce the time interval between the end of synthesis and the formation of the me-

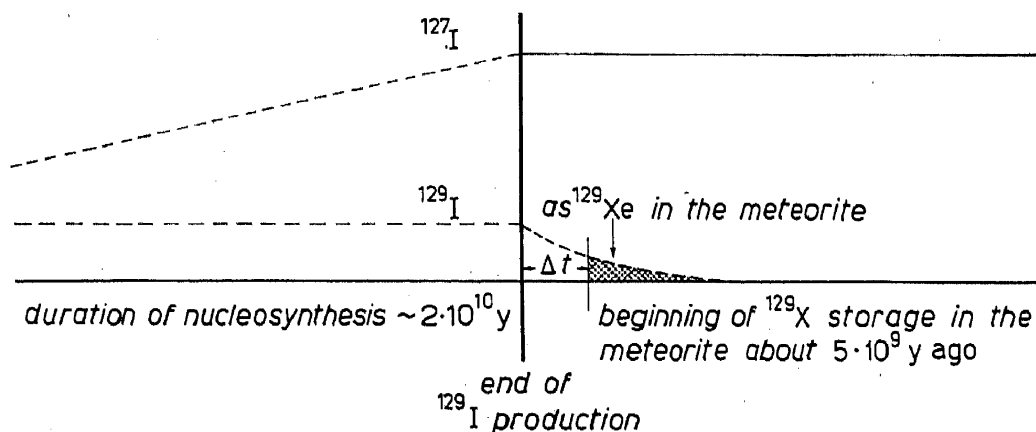


Fig. 4. — ^{129}I - ^{129}Xe content in dependence of time in the continuous synthesis model.

eteorite. In a sudden synthesis model the ^{127}I was about equal to ^{129}I . In a continuous synthesis model lasting about $2 \cdot 10^{10}$ years ^{129}I was at the end only 1% of the present ^{127}I content. Together with ^{127}I analysis of some chondrites [4, 5] REYNOLDS calculates for this time interval in the continuous synthesis model about $100 \cdot 10^6$ years. Figure 4 shows this I-Xe age in the meteoritic time scale.

The formula from which this time interval can be derived is

$$\Delta t = \frac{17.2}{0.69} [\ln (^{127}\text{I}/^{129}\text{Xe})_{\text{met}} + \ln (^{129}\text{I}/^{127}\text{I})_0] \text{ in } 10^6 \text{ years}$$

where $(^{129}\text{I}/^{127}\text{I})_0$ is the ratio at the end of the nucleosynthesis.

As mentioned above, FOWLER *et al.* [19] consider this time interval too

short and rather prefer the ^{129}I production in an intermediate stage of the solar system itself. In their model the planitesimals were destroyed before ^{129}I competely decayed and were then accumulated into larger bodies. The time needed for these parent bodies to condense from irradiated planitesimals lies between 10^6 and 10^7 years. These parents bodies (or body) were plane-toids of considerable size, in which a core and a mantle were differentiated. The differentiation and loss of the noble gas had to end in less than 10^8 years. The Δt as derived above, is in this model the time interval between the end of the irradiation and solidification of the meteorites. The ratio $^{129}\text{I}/^{127}\text{I}$ depends on the number of high-energy particles and on the amount of irradiated material. Again the short time scale is a problem in the formation of solid planets and their chemical differentiation.

The author has been tempted through experimental observation in studying the rare gas content of a large number of meteorites to favor the explanation that also ^{129}Xe has been largely trapped together with primordial Xe. In this model the following difficulties could be avoided:

1) ^{129}Xe has been so far always observed in company with primordial Xe, Kr and A.

2) The variations in the ^{129}Xe content are not so large as one would expect from the spread of ages in the formation of meteorites. If a spread of at least $200 \cdot 10^6$ years is assumed, as seems to be reasonable for larger bodies, the ^{129}Xe content should vary by a factor of 1000 and this has not been detected.

3) The time interval of about $100 \cdot 10^6$ years for the formation of the present meeorites out of protoplanetary material or of planitesimals is very short. The polimict structure indicates a complex history of meteoritic material.

To strengthen this hypothesis, a number of experiments has been performed which are described in the following.

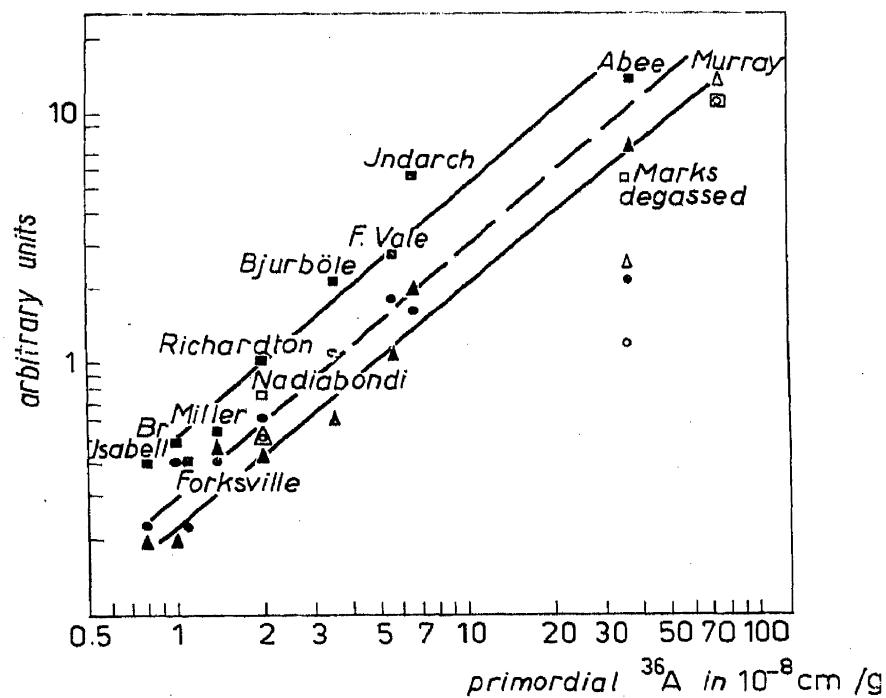
Seven samples have been selected out of 40 chondrites, which have a high $^{36}\text{A}/^{38}\text{A}$ ratio and indicate to contain some primordial argon. From spallation this ratio is expected to be about 0.6 and many chondrites have values as high as 5. The high ^{36}A content could also be explained by n-capture in ^{35}Cl as mentioned by GEISS and HESS [14]. From our experiments however, one has to conclude that the ^{36}A excess is of primordial origin, since all seven samples also contain Kr and Xe. All these chondrites also contain ^{129}Xe excess. The amount of the primordial gases shows some relation, which is seen in Fig. 5 where the ^{132}Xe , ^{129}Xe and ^{84}Kr values are plotted against primordial ^{36}A . The primordial ^{36}A content has been calculated from Table IV with the assumption that the $^{36}\text{A}/^{38}\text{A}$ from spallation is 0.6. Also the result of MURRAY [4] has been included and within the errors, which are

still large, the increase of these gases with primordial ^{36}Ar seems to establish this relation. St. Marks is an exception and has lower Kr and Xe values. Its K-A-age is also low and some degassing must have occurred. Anyway deviations like this are expected, since the mechanism of trapping primordial gases is not known. Very likely adsorption effects are involved and the trapped Ar-Kr-Xe -ratios might vary with temperature.

From this relation alone one cannot be certain that all the ^{129}Xe is trapped with the other primordial Xe isotopes. The primordial Xe and ^{129}I may have come into the sample in a very similar way and also ^{129}I can definitely be considered as primordial matter.

Therefore another kind of experiment is under progress [21]. It has been found that powdered samples release characteristic isotopes at different temperatures and that to a certain degree radiogenic spallation and primordial gases can be separated. In the achondrites we have seen that the primordial component escapes at lower temperatures than spallation products.

Fig. 5. — The ^{129}Xe , ^{132}Xe and ^{84}Kr content plotted vs. primordial ^{36}Ar : ■ □ ^{129}Xe , ● ○ ^{132}Xe , ▲ △ ^{84}Kr .



In the cases of St. Marks and Abee, as seen from $^{40}\text{Ar}/^{36}\text{Ar}$ ratio in Fig. 6 and 7, the radiogenic ^{40}Ar is released at lower temperatures than primordial ^{36}Ar . ^{36}Ar and also the xenon isotopes are expelled mainly at about 1200 °C in both cases. Only in the case of Abee at about 700 °C a small fraction of all primordial gases and ^{129}Xe escapes. At this temperature also a sulphur and a Ca deposit was observed and this is probably caused by the decomposition of oldhamite which is a mineral constituent of this meteorite [22]. The interpretation of the $^{129}\text{Xe}/^{132}\text{Xe}$ ratio is the following: Whenever an appreciable amount of Xe is released, the $^{129}\text{Xe}/^{132}\text{Xe}$ ratio is close to the value found in the analysis where a predegassed sample was totally melted. Only at lower temperature when atmospheric contamination from the sample and only a small amount of meteoritic xenon is released this ratio is lower. In this way the slow increase in the case of St. Marks can be understood.

The experiment on Abee has been repeated by REYNOLDS [15] who con-

firms the general feature of our result. Unfortunately he used fragments in his sample and those do always contain high atmospheric contamination. This has been clarified in several experiments by the author. In Reynold's

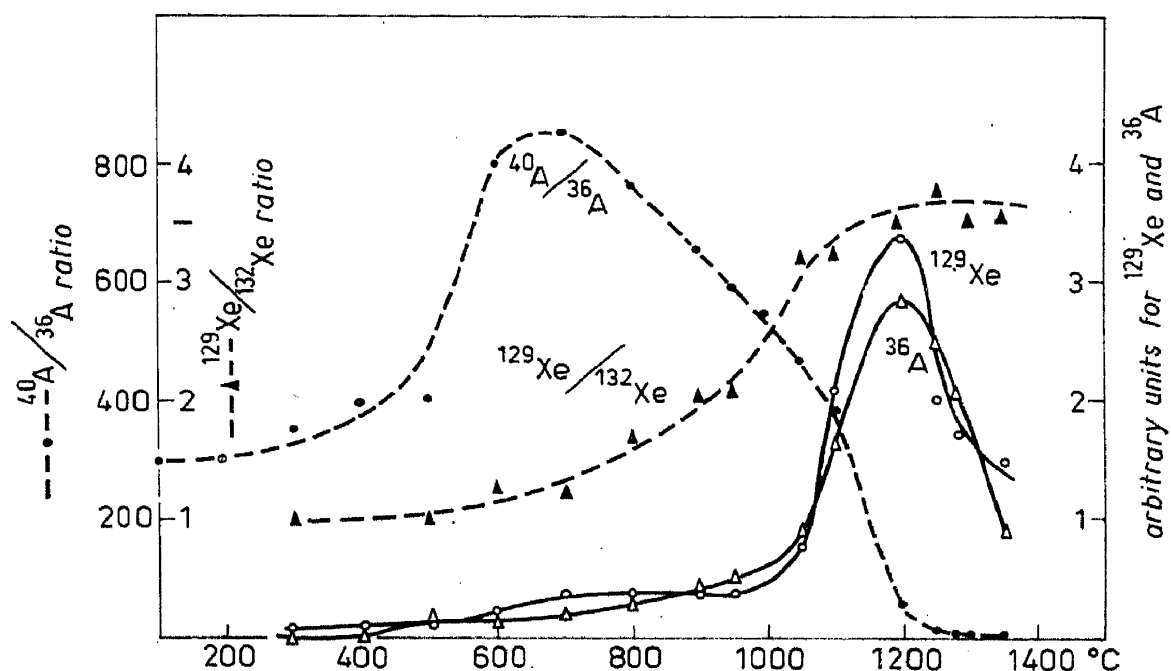


Fig. 6. — Degassing-experiment with the St. Marks enstatite chondrite.

experiment the $^{129}\text{Xe}/^{132}\text{Xe}$ ratio therefore varies more than in our experiment and again it is low when the amount of ^{129}Xe released is small. The main result which has to be drawn from these experiments shows in any case a strong connection of ^{129}Xe with primordial xenon and argon. In addition it shows

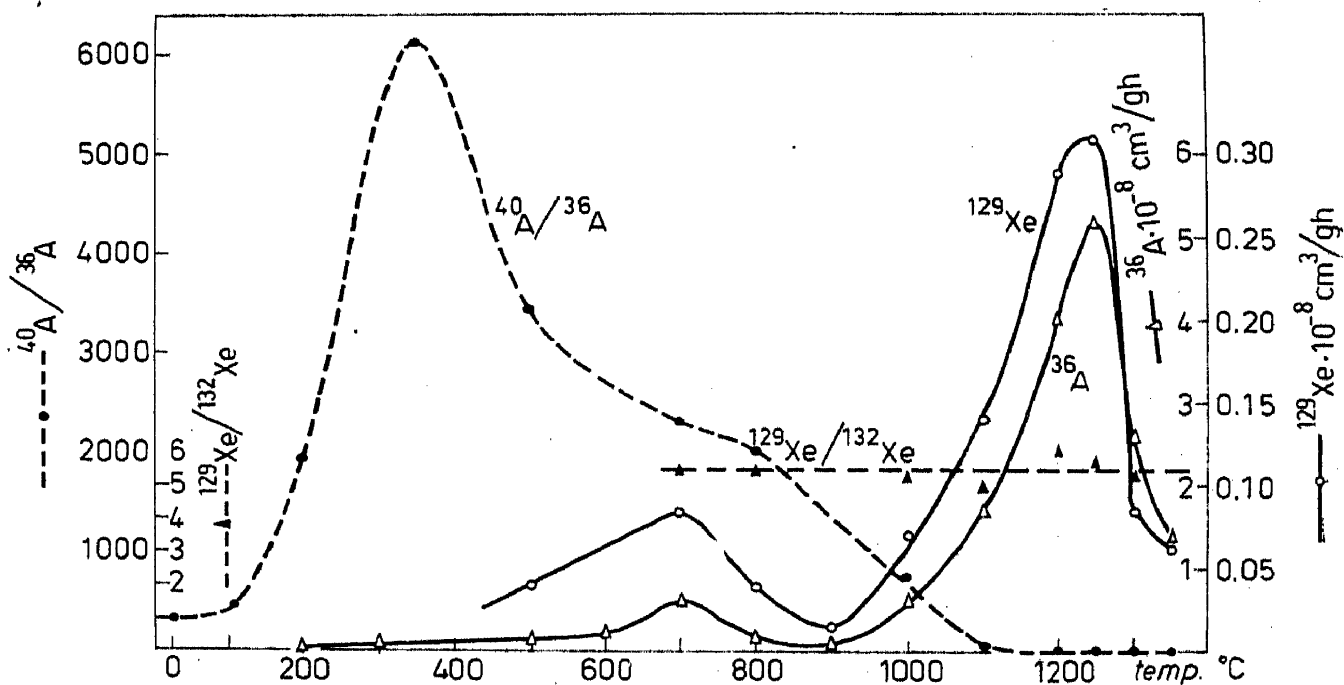


Fig. 7. — Degassing experiment with the Abee enstatite chondrite.

that ^{40}A is chemically differently bound and the K-A-age can only represent a lower limit. Thus an I-Xe-age cannot be added to the K-A-age in order to get the end of ^{129}I synthesis. This also explains the fact that meteorites such as St. Marks can have a very low K-A-age and still contain appreciable quantities of primordial gases. A relation between the ages and the primordial gas contents seems not to exist for these reasons. The author is also inclined to believe that earlier heating has influenced these meteorites and that only those components which are resistant to higher temperature could store the primordial gas. Such thermal treatment may also have acted as a mixing factor between radiogenic ^{129}Xe and the primordial Xe isotopes. With these remarks it should be indicated that any interpretation of the present results seems to be incomplete.

As mentioned above, this new field of meteorite research has only been in real progress for one year and the aim of this article has been not to give final answers to these far reaching problems but rather to show up the difficulties involved.

REFERENCES

- [1] O. F. SCHAEFFER and J. ZÄHRINGER: *Geochim. et Cosmochim. Acta*, **19**, 94 (1960).
- [2] E. K. GERLING and L. K. LEVSKII: *Dokl. Akad. Nauk. S.S.S.R.*, **101**, 750 (1956).
- [3] J. ZÄHRINGER and W. GENTNER: *Zeits. f. Naturfors.*, **15 a**, 600 (1960).
- [4] J. H. REYNOLDS: *Phys. Rev. Lett.*, **4**, 8, 351 (1960); *Journ. Geophys. Res.*, **65**, 3843 (1960); *Zeits. f. Naturfors.*, **15 a**, 1112 (1960).
- [5] E. ANDERS and C. M. J. STEVENS: *Journ. Geophys. Res.*, **65**, 3043 (1960).
- [6] V. R. MURTHY: *Phys. Rev. Lett.*, **5**, 539 (1960).
- [7] H. C. UREY: *Proc. Nat. Acad. Sci.*, **41**, 127 (1955).
- [8] I. KOHMAN: *Proc. N. Y. Acad. Sci.*, **62**, 503 (1956).
- [9] R. A. FISH, G. G. GOLES and E. ANDERS: *Astrophys. Journ.*, **132**, 243 (1960).
- [10] J. L. GREENSTEIN: *Astrophys. Journ.*, **113**, 531 (1951).
- [11] O. C. WILSON: *Publ. Astron. Soc. Pac.*, **73**, 15 (1961).
- [12] H. SUESS: *Journ. Geol.*, **57**, 600 (1949).
- [13] H. STAUFFER: *Geochim. et Cosmochim. Acta*, **24**, 70 (1961).
- [14] J. GEISS and D. C. HESS: *Astrophys. Journ.*, **127**, 224 (1958).
- [15] P. M. JEFFERY and J. H. REYNOLDS: *Zeits. f. Naturfors.*, **16a**, 431 (1961).
- [16] P. K. KURODA: *Nature*, **187**, 36 (1960).
- [17] A. G. W. CAMERON: *Icarus*, **1**, 13 (1962).
- [18] O. KRUMMENACHER, C. M. MERRIHUE, R. O. PEPIN and J. H. REYNOLDS: *Geochim. et Cosmochim. Acta*, **26**, 231 (1962).
- [19] W. A. FOWLER, J. L. GREENSTEIN and F. HOYLE: *Geophys. Journ.*, **6**, 148 (1962).
- [20] W. K. BONSAK and J. L. GREENSTEIN: *Astrophys. Journ.*, **131**, 83 (1960).
- [21] J. ZÄHRINGER and W. GENTNER: *Zeits. f. Naturfors.*, **16 a**, 239 (1961).
- [22] K. R. DAWSON, J. A. MAXWELL and D. E. PARSONS: *Geochim. et Cosmochim. Acta* **21**, 127 (196).

On the Origin of Cosmic Rays.

F. HOYLE

University of Cambridge - Cambridge

In the first part of this paper I shall attempt to review the orthodox point of view concerning cosmic rays arriving in the solar system from outside. Then I will discuss less orthodox considerations in a short second part. Finally, in the third part, the explicit problem of the origin of solar cosmic rays will be considered.

1. - Spatial distribution.

It will be assumed that external cosmic rays are of galactic origin, in particular that they originate in the galactic disk or perhaps in the galactic nucleus. It will be assumed that cosmic rays escape into the galactic halo, and that the whole system of the Galaxy is effectively closed—the escape of cosmic rays into extragalactic space is unimportant, except possibly at rigidities $> \sim 10^5$ GeV. In this connection it may be noted that evidence concerning the exposure of meteorites to cosmic rays, over a time interval long compared to the orbital period of the solar system in the Galaxy, precludes the possibility that cosmic rays are confined to a limited region of the galactic disk, while the radio-frequency emission from the halo of the Galaxy (and from the halos of other nearby galaxies) provides strong evidence of the existence of cosmic rays in the halo.

It may be accepted that the radiofrequency emission of the halo arises from synchrotron radiation by cosmic ray electrons. It cannot at present be definitely decided whether these electrons are of primary or secondary origin. If the electrons are secondaries, then the radio data leads to interesting and far-reaching conclusions concerning the primaries. These conclusions will be considered below. In view of the importance of the radio data to cosmic ray physics in general, it is clearly essential to establish whether the electrons are secondaries or not. If they are, then cosmic ray electrons arriving in the

solar system from outside would consist of a more or less equal mixture of negative and positive particles, whereas a large excess of negatives would be expected if the electrons are primaries.

In spite of the comparatively high abundance of heavy nuclei in the cosmic rays, the main contribution to the production of secondaries comes from collisions of the cosmic ray protons with protons of the ambient gas. The cross-section σ for such collisions is essentially independent of energy and is about $2 \cdot 10^{-26} \text{ cm}^2$. The number of protons per cm^3 with kinetic energies between E and $E+dE$, with E in GeV, is

$$(1) \quad \sim 5 \cdot 10^{-10} \frac{dE}{(1+E)^{\frac{3}{2}}}.$$

We examine the effect of this proton distribution, measured for the vicinity of the Earth, existing in the galactic halo.

The multiplicity s of electrons and positrons produced in a nuclear collision varies slowly with E . Over the limited range of E of interest in the radio problem we can take s to be constant, $s \simeq 10$. The total energy made available to the electron-positron secondaries $\sim 0.05 E$, so that the mean electron energy γmc^2 is given by

$$(2) \quad \gamma \simeq 10 E.$$

In the following work this mean energy will be used throughout. Strictly, account should be taken of the energy spectrum of the secondaries. However, it can be shown that such a refinement has little effect on the radio calculations (provided the shape of the spectrum does not change markedly with E).

The probability of a nuclear collision of a cosmic ray proton per unit time $\sim n\sigma \simeq 6 \cdot 10^{-16} \text{ s}^{-1}$. Multiplying by (1), and by $s \simeq 10$ gives the rate of production of secondaries by protons in the energy range E to $E+dE$. Since we shall be concerned with $E \gg 1 \text{ GeV}$, the denominator of (1) can be written as $E^{\frac{3}{2}}$, in which case

$$(3) \quad \sim 3 \cdot 10^{-24} n \frac{dE}{E^{\frac{3}{2}}} \text{ cm}^{-3} \text{ s}^{-1},$$

is the required production rate. This can be expressed in terms of γ , $d\gamma$ using (2). Thus

$$(4) \quad \sim 10^{-22} n \frac{d\gamma}{\gamma^{\frac{3}{2}}} \text{ cm}^{-3} \text{ s}^{-1},$$

is the rate of production of electrons with energies between γ and $\gamma+d\gamma$.

The extent to which the reservoir of electrons builds up depends on the

time scale over which the build up takes place. The time scale is to be taken as the least of the following four intervals:

i) The age of the Galaxy, about $4 \cdot 10^{17}$ s, or if cosmic rays have not always existed within the Galaxy a correspondingly shorter period.

ii) The time scale over which electrons lose an appreciable fraction of their energies by bremsstrahlung. This is $\sim 10^{26} n^{-1} \sigma^{-1}$ s, where $\sim 10^{-26}$ cm² is the bremsstrahlung cross-section for electrons with energies, γmc^2 , large compared to the rest mass ($\gamma \gg 1$).

iii) The time scale over which electrons lose an appreciable fraction of their energies by atomic collisions. If the ambient gas were neutral, this would be $\sim 2 \cdot 10^{12} \gamma n^{-1}$ s. Since the halo gas is mainly ionized this time scale is appreciable greater than $2 \cdot 10^{12} \gamma n^{-1}$ s.

iv) The time scale over which electrons lose an appreciable fraction of their energies by radio emission. This is $5 \cdot 10^8 \gamma^{-1} H^{-1}$ s, where H is the magnetic intensity.

Assuming that cosmic rays protons have always existed in the Galaxy, with the distribution (1), the expected secondary electron-positron reservoir is therefore given by

$$(5) \quad \sim 10^{-22} n \frac{d\gamma}{\gamma^{\frac{1}{2}}} \min [4 \cdot 10^{17}, 3 \cdot 10^{15} n^{-1}, > 2 \cdot 10^{12} \gamma n^{-1}, 5 \cdot 10^8 \gamma^{-1} H^{-2}].$$

Since we shall be concerned with $\gamma > 10^3$ the third of the above time scales is greater than the second, and can therefore be omitted from (5).

The rate of energy loss, $P(\gamma)$, through synchrotron radiation is given by

$$(6) \quad P(\gamma) \equiv \sim 1.6 \cdot 10^{-15} H^2 \gamma^2 \text{ erg s}^{-1}.$$

Provided we are not concerned with radiation near the ends of an electron distribution such as (5), it is sufficient to take the emission as occurring at a particular frequency ν , given by

$$(7) \quad \nu = 4.2 \cdot 10^6 H \gamma^2 \text{ Hz}.$$

Strictly, the magnetic intensity H should be replaced both in (6) and (7) by the component of the magnetic field perpendicular to the electron velocity, but this refinement can also be omitted.

The emissivity, and the radio spectrum, is obtained by multiplying (5) and (6), and by replacing γ , $d\gamma$ by ν , $d\nu$ with the aid of (7). The result is

$$(8) \quad \sim 2 \cdot 10^{-39} \frac{d\nu}{\nu^{0.75}} n H^2 \min (4 \cdot 10^{17}, 3 \cdot 10^{15} n^{-1}, 10^{12} \nu^{-\frac{1}{2}} H^{-\frac{1}{2}})$$

If the third alternative in the brackets were the least, the radio spectrum would take the form $d\nu/\nu^{1.25}$. Since the observed spectrum is certainly not at this, we can infer that

$$(9) \quad \nu^{-\frac{1}{2}} H^{-\frac{3}{2}} > \min (4 \cdot 10^5, 3 \cdot 10^8 n^{-1}),$$

for the values of ν concerned in the observation data. A reasonable choice for ν would be $\nu = 1.6 \cdot 10^8$ Hz. Inserting this value in (9) gives an upper limit for the magnetic intensity in the halo

$$(10) \quad H < \sim 10^{-6} [\min (5, 4 \cdot 10^{-2} n^{-1})]^{-\frac{2}{3}} \text{ gauss.}$$

Subject to the assumptions made above, we see that the magnetic intensity in the halo cannot exceed $\sim 3 \cdot 10^{-7}$ gauss, unless n exceeds $\sim 10^{-2}$ atom cm^{-3} . Reducing the time for which cosmic rays have existed in the halo would increase the permitted upper limit, as would an appreciable escape of electrons into extragalactic space.

It is of interest to consider the emissivity on the basis that H takes its maximum permitted value. It is then easy to show that the radio emission is determined by

$$(11) \quad \sim 3 \cdot 10^{-33} n \frac{d\nu}{\nu^{0.75}}.$$

This may be compared with the observed emissivity of the galactic halo and of other normal Sb and Sc spirals, which at $\nu = 1.6 \cdot 10^8$ Hz are of order $3 \cdot 10^{-41}$ erg $\text{cm}^{-3} \text{ s}^{-1} (\text{Hz})^{-1}$. At this frequency, (11) gives

$$(12) \quad \sim 2 \cdot 10^{-39} n \text{ erg cm}^{-3} \text{ s}^{-1} (\text{Hz})^{-1}.$$

Hence we obtain a remarkable agreement with observation for $n \simeq 10^{-2}$ atom cm^{-3} .

Because of the far reaching implications of this conclusion, it will be as well to give an independent argument leading to a similar value for n . Even if H is set as high as 10^{-6} G, the magnetic field by itself cannot restrain the cosmic rays from flying off into extragalactic space, if the cosmic rays possess an energy density comparable to that found in the vicinity of the earth. For the magnetic field to exercise an appreciable restraining influence it is necessary that $H^2/8\pi$ be at least 1 eV cm^{-3} , i.e. that $H \sim 5 \cdot 10^{-6}$ G. Such a high intensity seems entirely inconsistent with the radio frequency spectrum. (This remark lies within the framework of the present point of view concerning cosmic rays in the Galaxy. It does not apply to the unorthodox development of Section 2). To restrain the cosmic rays, a reservoir of ambient gas is necessary, with a value of n high enough for the gravitational potential energy per

unit volume to exceed $\sim 1 \text{ eV cm}^{-3}$. Since the gravitational potential energy in the halo $\sim 5 \cdot 10^{-10} \text{ erg atom}^{-1}$, we hence require $5 \cdot 10^{-10} n > 1.6 \cdot 10^{-12}$, *i.e.* $n > \sim 3 \cdot 10^{-3} \text{ atom cm}^{-3}$.

With $n \simeq 10^{-2} \text{ atom cm}^{-3}$ in the halo, a cosmic ray proton has a probability $\sim 6 \cdot 10^{-18} \text{ s}^{-1}$ of experiencing a nuclear collision. Therefore in a time scale $\sim 10^{17} \text{ s}$ an appreciable fraction of protons experience a nuclear collision. Indeed the cosmic ray protons have a lifetime comparable with, but less than, the age of the Galaxy. The heavy nuclear component of the cosmic rays, possesses still shorter lifetimes. On the other hand the chemical composition of cosmic rays in the vicinity of the Earth shows very clearly that the nuclear component with cannot have traversed more than $\sim 3 \text{ g}$ of hydrogen. A paradox evidently arises. The following suggestions can be made for overcoming it.

a) The cosmic ray protons and the heavy nuclear component originate in different processes. If the Earth happens to lie near a source of the heavy component, the intensity of the latter could be abnormally high, and the amount of hydrogen traversed could be correspondingly low. This suggestion is implausible, however, in view of the similarity of the spectra of protons and heavy nuclei up to rigidities of order $\sim 10 \text{ GeV}$.

b) The electrons responsible for the radio emission are primaries. The radio emission argument leading to the low value of H in the halo and to the high values of n would then be inapplicable. Thus n might be set as low as $\sim 10^{-4} \text{ atom cm}^{-3}$ and the amount of hydrogen traversed in the age of the Galaxy would be no more than $\sim 5 \text{ g}$ for a cosmic ray within the halo. However, it is to be expected that cosmic rays spend a fraction of the time in the disk of the Galaxy. Taking this fraction equal to the ratio of the volume of the disk, $\sim 10^{66} \text{ cm}^3$, to the volume of the halo $\sim 10^{69} \text{ cm}^3$, the time spent in the disk $\sim 3 \cdot 10^{14} \text{ s}$. With a density of $\sim 1 \text{ atom cm}^{-3}$ in the disk, the mass of hydrogen traversed $\sim 15 \text{ g}$. Hence, unless cosmic rays in the halo cannot reach the disk, the paradox survives. It is difficult to see why any such prohibition should exist, since cosmic rays originating in the disk can be expected to return eventually to it. Moreover, the agreement of the emissivity calculated above from (12) with the observed emissivity would be fortuitous, which seems rather unlikely.

c) A more plausible resolution can be obtained along the following lines. Write V_h , V_d for the respective volumes of the halo and the disk and T for the average time spent by a cosmic ray particle in the disk, before it passes into the halo. If initially there were no cosmic rays in the halo, a time TV_h/V_d would be required for the energy density in the halo to rise to a value similar to that in the disk. At this stage approximately as much cosmic ray

energy would re-enter the disk as left it. As time went on, the energy density, both in the disk and the halo, would continue to rise, provided there was no significant absorption by nuclear collisions; and a growing fraction of the cosmic rays in the disk at any moment would be particles that had spent a time interval of at least TV_h/V_d in the halo. (In the latter connection it may be noted that, since H is probably appreciably lower in the halo than it is in the disk, particles do not readily return to the disk. But as equilibrium tends to be set up, with the rising energy density, particles must return to the disk, to the extent required by the quasi-equilibrium.)

An upper limit can be set to T , from the condition that the nuclear component of the cosmic rays with $Z \geq 3$ has not passed through more than ~ 3 g of hydrogen. This corresponds to $\sim 2 \cdot 10^{24}$ atom cm^{-2} . For a disk density of about 1 atom cm^{-3} , the maximum path length is therefore $\sim 2 \cdot 10^{24}$ cm, which is traversed by a cosmic ray in $\sim 7 \cdot 10^{13}$ s. Hence T must be $\leq 7 \cdot 10^{13}$ s. The interesting case is that of equality, $T \simeq 7 \cdot 10^{13}$ s $\cong 2 \cdot 10^6$ years.

With $V_h/V_d \simeq 10^3$, the average time spent by a halo cosmic ray before returning to the disk is then $\sim 7 \cdot 10^{16}$ s. The path traversed is $\sim 2 \cdot 10^{27}$ cm. For $n \simeq 10^{-2}$ atom cm^{-3} , the particle therefore traverses ~ 30 g of hydrogen. Since $\sigma \simeq 2 \cdot 10^{-26}$ cm^2 for nuclear collisions by protons, the chance of a proton experiencing a nuclear collision in traversing this quantity of hydrogen is $1 - \exp[-0.4] \simeq \frac{1}{3}$. Hence a time TV_h/V_d can elapse without more than a modest fraction of the halo protons experiencing collision. This implies that the energy densities in the halo and disk will indeed be comparable, as was assumed in the above calculation of the radio emission from the halo. Moreover, we expect that, since the time scale for nuclear collision is comparable with TV_h/V_d , roughly one half of the energy density in the disk at any moment will consist of cosmic rays that have returned from the halo, and which have traversed ~ 30 g of hydrogen. The other half consists of cosmic rays that have not yet entered the halo, that have originated during the last $2 \cdot 10^6$ years, and which have traversed only ~ 3 g of hydrogen. It follows that the cosmic rays divide into two groups, and that the energy contribution of the two groups are approximately equal.

The apparent paradox is now resolved, for only the group that has traversed ~ 3 g of hydrogen can contain heavy nuclei. Cosmic rays returning to the disk from the halo must consist essentially of protons, D, ^3He , and ^4He , since in traversing ~ 30 g of hydrogen all nuclei from lithium upwards would be destroyed. Although at a given rigidity protons are in great excess by number, at a given energy their contribution is closely comparable to that of nuclei with $Z \geq 6$. This approximate energy equality could well correspond to the approximate equality of our two groups. Hence we can follow a suggestion of PETERS, that *all* cosmic rays might start as heavy nuclei. Hitherto the difficulty in accepting this attractive idea has been the low value of the

cosmic ray ratio of (Li, Be, B) to (C, N, O). This is not consistent with a *well-mixed* assembly in which heavy nuclei, Fe say, are steadily degraded to protons. The difficulty does not arise here, however, since the cosmic rays are not well mixed in the present model. An observer far out in the halo would find (Li, Be, B) to be more abundant than (C, N, O).

Two tests of these ideas follows immediately. The isotopes ^3He and ^4He must be comparable in their abundances. According to the recent work of APPA RAO [1] this appears to be the case for rigidities less than 1.6 GeV. At such low rigidities the ^3He could be of solar origin, however. Thus, although a failure to find ^3He would contradict the above ideas, detection at low rigidities cannot yet be taken as confirming them.

Electrons enter the disk from the halo. If, as seems reasonable, the rate of entry is comparable to that of protons of the same rigidity, we can expect the electron density (say for $\gamma > 2000$) to attain equilibrium with the halo reservoir in a time scale $T \simeq 7 \cdot 10^{13}$ s. (Collision with the neutral hydrogen atoms in the disk do not degrade the electron energies at all seriously. The time scale for degradation at a density of about 1 neutral atom cm^{-3} is $2 \cdot 10^{12} \gamma^{-1}$ s, which for $\gamma > 2000$ is long compared to T .) Hence the electron density in the disk should be similar to the distribution (5), where n is the halo density. With $n \simeq 10^{-2}$ atom cm^{-3} , $H \simeq 3 \cdot 10^{-7}$ G, the distribution for $2 \cdot 10^3 < \gamma < 2 \cdot 10^4$ is

$$(13) \quad \sim 3 \cdot 10^{-7} \frac{d\gamma}{\gamma^{\frac{3}{2}}} \text{ cm}^{-3}.$$

To obtain the expected flux of electrons more energetic than γ , for γ values not too close to the upper limit of applicability of (13), first integrate with respect to γ (giving $\sim 2 \cdot 10^{-7}/\gamma^{\frac{1}{2}}$) and then multiply by $10^4 c/4\pi$ in order to obtain a value expressed in $\text{m}^{-2} \text{s}^{-1} \text{sr}^{-1}$. The result is $\sim 5 \cdot 10^6/\gamma^{\frac{1}{2}} \text{ m}^{-2} \text{s}^{-1} \text{sr}^{-1}$. With the electron energy in GeV, the flux of electrons more energetic than E is $\sim 60/E^{\frac{1}{2}} \text{ m}^{-2} \text{s}^{-1} \text{sr}^{-1}$. For comparison the proton flux obtained from (1) is $\sim 10^4/(1+E)^{\frac{1}{2}} \text{ m}^{-2} \text{s}^{-1} \text{sr}^{-1}$.

At $\gamma > 2 \cdot 10^4$ ($E > 10$ GeV) the distribution (13) is not applicable, since degradation by synchrotron radiation provides the dominating time scale for electrons in the halo. The corresponding distribution

$$(14) \quad \sim 5 \cdot 10^{-14} n H^{-2} \frac{d\gamma}{\gamma^{\frac{3}{2}}} \simeq 5 \cdot 10^{-3} \frac{d\gamma}{\gamma^{\frac{3}{2}}},$$

for $n \simeq 10^{-2}$ atom cm^{-3} , $H \simeq 3 \cdot 10^{-7}$ G. Again expressing the electron energy in GeV, the electron flux for $E > 10$ GeV is $\sim 300 E^{-\frac{1}{2}} \text{ m}^{-2} \text{s}^{-1} \text{sr}^{-1}$. The ratio of the electron flux to the proton flux at $E > 10$ GeV is therefore $\sim 0.03 E^{-1}$.

Results may be summarized as follows:

E (GeV)	Electron flux/Proton flux (per cent)
1	~ 1.5
10	0.3
100	0.03
1000	0.003

These flux ratios are given for the same electron and proton kinetic energies. At rigidities of 10, 10^2 , 10^3 GeV the ratios are of course essentially the same. But at a rigidity of 1 GeV the ratio is reduced to about 1%. This value lies somewhat above the upper limit set by CRITCHFIELD, NEY and OLESKA [2]. The discrepancy does not seem a crucial one, however, since none of the above formulae can be taken as accurate within a factor 2.

Processes of origin. — At first sight it seems easier to understand how the cosmic rays might have acquired a definite energy spectrum if they originated in one source only, such as has been proposed by GINZBURG and PIKELNER, a source in the galactic nucleus. However, an argument against a single source can be given, although the numerical values involved are perhaps not sufficiently secure to make the argument decisive.

In terms of the model described above, it is necessary that the time scale for escape into the halo be $\sim T$, irrespective of the energy of a particle—otherwise the heavy nuclear component would not show the same spectrum as the protons and helium nuclei. This implies that the process is one of diffusion. Writing h for the thickness of the disk, the mean free path in the diffusion process $\sim 3h^2/cT$, which with $h \simeq 200$ parsec, $T \simeq 7 \cdot 10^{13}$ s, is about $\frac{1}{5}$ parsec. If the cosmic rays originate in a single process then this mean free path cannot be interpreted in an isotropic sense, otherwise particles would only diffuse a distance h from the source in time T . The mean free path $\sim \frac{1}{5}$ parsec must apply only in a direction perpendicular to the galactic plane. In time T , particles must move a distance of at least the galactic radius r in direction parallel to the plane. Hence the mean free path in such directions must be some 10^4 times greater than in the normal direction. Indeed, if particles travel along the spiral arms, the distance they must move in order to fill the galactic disk is considerably greater than r , being $\sim 2\pi r$ multiplied by the number of turns of the spiral arms, which probably ~ 5 . Taking $r \simeq 5$ kpc, this is $\sim 5 \cdot 10^{23}$ cm. Even this distance is probably too small, since spiralling around the magnetic field lines increases the path length, say by a further factor ~ 3 , which yields a total of $\sim 1.5 \cdot 10^{24}$ cm, requiring a travel time close to T . Hence it seems that, while diffusion with path length

$\sim 10^{18}$ cm takes place perpendicular to the galactic plane, there must be an essentially simple streaming of cosmic rays in directions parallel to the plane. Such a streaming implies a lack of isotropy. In time T the cosmic rays produced by a single source have scarcely sufficient time to fill the disk before they escape into the halo. This requires the flow in the disk to possess a directed sense.

It can be argued that the observed isotropy of cosmic rays is a product of deflections taking place locally within the solar system, but this suggestion does not seem very plausible at energies above 10^3 GeV. A lack of isotropy, not in rigidity experiments (at a given rigidity most cosmic rays are protons, and these, coming from the halo in the above model would surely be isotropic), but in air shower experiments, is to be expected. No appreciable anisotropy is in fact found. The argument appears therefore to point against a single source, although the circumstance that the time interval required to fill the disk by simple streaming turns out to be of order T might be taken to support the idea of a single source.

A stellar origin of cosmic rays does not suffer from any such difficulty, since many sources would be operative over a time interval $\sim T$. In $\sim 2 \cdot 10^6$ years some 10^4 supernovae might be expected. If they occur at random over the galactic disk, the average area of the disk per supernova is $\sim 3 \cdot 10^{41}$ cm², which for a circular area corresponds to a diameter of ~ 200 parsec. Since this is precisely of the order of the width of the disk, the mean free path computed above can be interpreted in an isotropic sense.

The supernovae supply an adequate total amount of energy. Taking the volume of the galactic disk as $\sim 3 \cdot 10^{66}$ cm³, the total cosmic ray content at 1 eV cm^{-3} is $\sim 5 \cdot 10^{54}$ erg. If this is to be provided by $\sim 10^4$ supernovae the requirement is for $\sim 5 \cdot 10^{50}$ erg per supernova. FOWLER and HOYLE [3] give the energy output of type II supernovae as greater than 10^{51} erg. The energy requirement is tight but not impossible. Since it also seems likely that the type II supernovae are the seat of origin of the iron group elements [3] such a source fits the chemical requirement very well indeed.

Moreover, intense electromagnetic activity is certainly associated with expanding supernova envelopes. In the spectacular case of the *Crab Nebula* $\sim 10^{48}$ erg are present as cosmic ray electrons. If protons are also present, the total cosmic ray energy could be higher still. Nor is the *Crab Nebula* an isolated case. Increasingly large numbers of non-thermal radio sources have been found recently along the plane of the Galaxy. These are very likely supernova remnants. Indeed it seems probable that a considerable fraction of the radio emission from the plane of the Galaxy comes from such remnants. Their number could well be of order 10^3 or more.

An investigation I made recently [4] on the cause of the electromagnetic activity in supernova remnants would, if correct, weaken the latter part of

this argument, however. It appeared that the rapidly expanding gaseous envelope acts on an already-existing external cosmic ray assembly in such a way as to lift the energies of the cosmic ray particles. The process is analogous to the Fermi mechanism in which collisions with massive clouds lift the cosmic-ray energies, but in which the cosmic rays must first be present. This would imply that supernova remnants are not the seat of origin of cosmic rays, although a considerable fraction of the energy of expansion of the envelope can be taken up in lifting the energies of the already-existing cosmic rays.

Nor does it seem likely that cosmic rays can originate at the moment of explosion of a supernova. The density of matter in, or near, a supernova is likely to be so great that heavy nuclei accelerated to relativistic energies would be almost instantly fragmented. The moment of origin must, it seems, occur at least some days after the first outburst. This can be seen by writing U for the expansion velocity. After time t the surface area of the expanding envelope is $4\pi U^2 t^2$. If we consider a layer 1 g thick at the surface of the envelope the mass is just $4\pi U^2 t^2$. If this mass should become accelerated to relativistic energy the total energy $\sim 4\pi U^2 t^2 c^2$. Equating to $\sim 10^{51}$ erg, and putting $U \simeq 3000 \text{ km s}^{-1}$, gives $t \simeq 10^6 \text{ s}$. This is the minimum time interval that must elapse between the first outburst and the acceleration of cosmic rays. Otherwise the cosmic rays would need to be produced in so thick a layer that a break up of heavy nuclei would ensue. A still greater time interval $t \sim 1$ year, say, would be desirable, in which case the envelope would have attained a radius $\sim 10^{16} \text{ cm}$, very much larger than the original star, which can scarcely have exceeded 10^{10} cm in radius at the time of explosion.

The question arises as to how such a delayed acceleration can arise. Hardly by any outbreak of turbulence. After an interval of 1 year the expansion of the envelope must surely have reduced very greatly any initial internal energy possessed by the outward moving material. A second phase of explosion of the star is needed, in which matter is ejected at still higher speed than in the first outburst. The matter from the second explosion then catches up with that from the first explosion. The effect is to produce an outward moving shock that travels through the original envelope until it reaches the surface. The initial speed of the shock is of the order of the speed of the second outburst. This could be very high indeed, $\sim 10^9 \text{ cm s}^{-1}$. As the shock expands into lower density regions, as it approaches the surface of the outer envelope, its speed must rise, and no very great increase is necessary to bring it into the relativistic range. It does not appear unreasonable to suppose that a surface layer of the outer envelope might be thrown off with a kinetic energy comparable to its rest mass.

This possibility has affinities with a suggestion of COLEGATE and JOYNSON. It differs in detail in the requirement for a second outburst, in not attributing the origin to the first explosion, and in not seeking necessarily to explain the

cosmic ray energy spectrum in terms of the behaviour of the shock as it approaches the outer surface. It may be sufficient simply for the process to be one of injection at the relativistic threshold, as will be seen in a moment.

Since it became clear that type II supernovae possess masses much larger than the sun [3], the eventual evolution of the stellar remnant remaining after the first outburst has raised interesting questions. The remnant cannot be disintegrated by the first outburst, nor is its mass at all likely to be below the degeneracy limit. A second outburst therefore seems certain to occur. Moreover since the second outburst occurs at a more compressed state than the first it must take place at higher speed. In the past, the difficulty has been to see how the second outburst could take place as soon as a year, or even 10^2 years, after the first. The remnant in the type II case is initially non degenerate—it possesses a considerable store of thermal energy. Not until sufficient time has elapsed for this energy to leak out of the star can a second drastic process take place. The time scale for leakage was formerly thought to be too long for any interesting overlap, of the sort considered above, to take place. But recent work on rapid neutrino emission at temperature $\sim 10^9$ °K suggests that the leakage time scale must be set shorter by many orders of magnitude than loss by radiative emission would give. Indeed, neutrino leakage could well reduce the time scale for evolution of the stellar remnant to ~ 1 year. It will be amusing, and rather ironic, if the origin of cosmic rays should turn out to depend quite crucially on the properties of the neutrino.

The double outburst of a supernova has the further advantage of permitting an explanation of the power law energy spectrum of cosmic rays, provided one is willing to relax somewhat the diffusion condition parallel to the galactic plane, provided motions parallel to the plane allow cosmic rays produced by one supernova to enter regions influenced by other supernovae. If in time T cosmic rays diffuse ~ 1 kpc parallel to the plane as compared with ~ 100 parsec normal to the plane, the necessary requirement is met.

The impressive feature of Fermi's theory was the derivation of a power law. The effect of collisions was to produce an exponential gain rate α ,

$$(15) \quad \frac{dE}{dt} = \alpha E .$$

Combined with this, it was assumed that the probability of a cosmic ray escaping from the region of acceleration was a constant, β say, per unit time, and that β was independent of the cosmic ray energy E . Then it was easy to show that the cosmic rays acquired the integral spectrum $E^{-\beta/\alpha}$.

These ideas can be incorporated into the supernova theory and the ratio β/α can be shown to be of order unity, subject to certain straightforward assumptions.

Supernovae cannot be expected in any strict sense to be uniformly distributed over the face of the Galaxy. Rather it is to be expected that in any particular interval of time supernovae are concentrated in discrete complexes of interstellar gas clouds. Let q be the supernova rate in one such complex. We consider, first, the cosmic rays in this one complex which we take to be of volume V_h . Let the probability of escape of a cosmic ray out of the complex into the general disk of the Galaxy be $\tau^{-1} \text{ s}^{-1}$, with $\tau \ll T$. Assume that within the complex cosmic rays are well mixed with a total energy W for the whole complex. Write K for the ultimate expansion energy of a supernova (after second outburst) and k for the energy of the cosmic rays generated at the second outburst.

Returning to the cosmic ray energy spectrum, we have $\beta = \tau^{-1}$. To obtain α we note that in a time interval dt there are $q dt$ supernovae. The dynamical energies possessed by their expanding envelopes $\sim Kq dt$. (The interval dt we take to be ~ 1 year, so that we are concerned with dynamical energies after second outburst.) This energy is used to lift the energies of already-existing cosmic rays. Hence W is increased by $kq dt$ because of newly generated cosmic rays and by $Kq dt$ because of the lifting of already-existing cosmic rays. On the other hand W is decreased by the escape of cosmic rays out of the complex, the amount of the decrease being $W\tau^{-1}dt$ in the time interval dt . On the basis that W reaches a more or less steady balance, we can equate gains and losses,

$$(16) \quad W = \tau q(k + K),$$

this being an equation for W . Furthermore, the energy E of a single cosmic ray particle increases to $E + dE$ in the time interval dt , where

$$(17) \quad \frac{dE}{E} = \frac{Kq}{W} dt,$$

provided all cosmic rays partake in the lifting process in proportion to their energies (this was the situation investigated in [4]). In this case $\alpha = Kq/W$ and, inserting the above expression for W , we have $\alpha = K\tau^{-1}(k+K)^{-1}$. With $\beta = \tau^{-1}$, the ratio β/α is $1 + k/K$, and the cosmic ray spectrum in the complex has integral power law. It is remarkable that the form is determined only by the ratio k/K . The quantities τ , q , W do not enter.

The next step is to consider the ratio k/K . Write M_1 , M_2 , for the masses ejected in the first and second outbursts, and let U_1 , U_2 be the corresponding velocities. After the second envelope catches the first, the outward speed becomes $(M_1 U_1 + M_2 U_2)(M_1 + M_2)^{-1}$, and the loss of dynamical energy is $\frac{1}{2} M_1 M_2 (U_2 - U_1)^2 (M_1 + M_2)^{-1}$. This is the energy available for the shock. Thus

we set

$$k \simeq \frac{1}{2} \frac{M_1 M_2}{M_1 + M_2} (U_2 - U_1)^2,$$

the remaining energy K being given by

$$K \simeq \frac{1}{2} \frac{(M_1 U_1 + M_2 U_2)^2}{M_1 + M_2}.$$

According to FOWLER and HOYLE, M_1 and M_2 are likely to be comparable, in which case

$$\frac{k}{K} \simeq \frac{(U_2 - U_1)^2}{(U_2 + U_1)^2}.$$

On this basis we expect an integral energy spectrum with exponent between 1 and 2, in agreement with observation.

So far we have considered only the situation within one particular complex of interstellar clouds in which type II supernovae are occurring. The question now arises as to what will be the energy spectrum of the general distribution of cosmic rays in the disk of the Galaxy. This distribution arises from superposition of cosmic rays from many separate complexes. But provided k/K is the same, or has the same range of variation, for each complex, and provided cosmic rays are generated at the same injection energy in each complex, the contribution of many complexes will combine into a distribution possessing the same energy spectrum as that of each separate complex. It is not unreasonable to suppose that these conditions are satisfied. All we require is that the supernovae be closely similar in the different complexes.

It remains to consider ordinary stable stars as sources of cosmic rays. Taking main-sequence stars first, O and B stars do not possess convection zones, and hence do not possess sources of free energy near their surface. Hence it seems unlikely that very early-type main-sequence stars can be major contributors of cosmic rays, in spite of their high individual luminosities. Dwarf stars possess deep sub-photospheric convection zones, and very likely electromagnetic activity of the kind observed at the solar surface occurs on all dwarf stars. The dwarf flare stars are notable cases. It must be remembered, however, that flares are conspicuous in dwarf flare stars only because the normal luminosities of these stars are very low. There is no reason to believe that the electromagnetic activity in dwarf stars is in general any greater than the solar activity. It seems then that the dwarf stars fall short of supplying cosmic rays at the required rate by many orders of magnitude.

The most likely sources of cosmic rays among main-sequence stars are the A stars. The A stars possess convection zones. They rotate much more rap-

idly than the dwarf, and the combined effects of convection and rotation could produce enhanced electromagnetic activity. Intense magnetic fields, $\sim (10^3 \div 10^4)$ G, have been observed at the surfaces of many A stars. Moreover, their surface chemical compositions are often peculiar in a way that points to pronounced electromagnetic activity—to nuclear reactions following acceleration to millions of volts at the least. Yet the total energy requirement appears too high for the A stars, the total luminous output of which can scarcely exceed $\sim 10^{42} \text{ erg s}^{-1}$ for the whole Galaxy. Unless these stars convert some 10% of their whole output into cosmic rays they fall short of meeting the energy requirement. While it is true that a similar measure of efficiency is necessary for the supernovae, the energy of supernovae is free energy, whereas only a small fraction of the energy of the A stars is free energy.

Most of the luminous output of the Galaxy comes from the giant stars, an output $\sim 3 \cdot 10^{43} \text{ erg s}^{-1}$. An efficiency of $\sim 0.1\%$ would make the giants important contributors of cosmic rays. Giants, moreover, possess highly active sub-photospheric convection zones, and hence considerable free energy is available near their surfaces. Little appears to be known about this interesting possibility, which should evidently be kept in mind, in spite of the apparent success of the supernova theory.

2. — The radio-spectrum difficulty.

The spectrum obtained in Section 1 for the halo of the Galaxy was $d\nu/\nu^{0.76}$. This is close to the form $d\nu/\nu^{0.7}$ quoted by some observers, but is distinctly steeper than the form $d\nu/\nu^{0.5}$ given by others. The difference is at first sight a minor one, but the implications of the difference are far reaching, so far reaching as to justify a more thoroughgoing observational investigation than has hitherto been made.

There appears to be little scope for changing the form of the calculated electron production spectrum. If account had been taken of the dependence of electron multiplicity s on the primary energy E , the effect would have been to steepen the spectrum very slightly, to about $d\nu/\nu^{0.8}$. This appears to be of the order of the effect of the approximations used in the calculation. Hence to flatten the spectrum to $d\nu/\nu^{0.5}$, if indeed the spectrum should be as flat as this, something other than a modification of this part of the calculation appears necessary.

To yield a radio spectrum $d\nu/\nu^{0.5}$ the electron spectrum must be of the form $d\gamma/\gamma^2$, instead of $d\gamma/\gamma^{\frac{3}{2}}$. Such a distribution arises under the following conditions. Suppose electrons are produced only for $\gamma > \gamma_1$, and let the magnetic intensity be high enough for degradation through synchrotron radiation to take place down to $\gamma_0 \ll \gamma_1$. Furthermore, let a steady-state be set up in

which the rate of production of secondaries equals the rate of degradation. That is to say, the rate at which electrons are degraded below γ ($\gamma_0 < \gamma < \gamma_1$) is equal to the rate of electron production. Taking this to be constant, the rate of degradation below γ is a constant—the number of electrons above γ at time t that have fallen below at time $t+dt$ is just the number of electrons produced in this time interval. Write $N(\gamma)d\gamma$ for the electron distribution per unit volume in this steady state. Then the number of electrons degraded below γ ($\gamma_0 < \gamma < \gamma_1$) is simply $N(\gamma)P(\gamma)/mc^2$ per unit volume per unit time. If this is to be equal the production rate, then

$$(18) \quad N(\gamma) = \frac{mc^2}{P(\gamma)} \quad (\text{electron production rate}).$$

Inserting $P(\gamma)$ from (6) gives $N(\gamma) \propto \gamma^{-2}$.

The total emission per unit volume from electrons between γ and $\gamma+d\gamma$ is just $N(\gamma)P(\gamma)d\gamma$. Using (18) together with $d\gamma \simeq 2.5 \cdot 10^{-4} H^{-\frac{1}{2}} d\nu/\nu^{0.5}$ from (7), gives an emission rate

$$(19) \quad 2.5 \cdot 10^{-4} \cdot H^{-\frac{1}{2}} \cdot mc^2 \cdot \frac{d\nu}{\nu^{0.5}} \quad (\text{electron production rate}).$$

Since we require the emissivity at $\nu \simeq 10^8$ Hz to lie between 10^{-40} and 10^{-41} erg cm $^{-3}$ s $^{-1}$ /Hz, it follows that the electron production must satisfy the condition

$$(20) \quad (\text{electron production rate}) \simeq 10^{-33} H^{\frac{1}{2}}/mc^2 \simeq 10^{-27} H^{\frac{1}{2}} \text{ cm}^{-3} \text{ s}^{-1},$$

H being in gauss.

It is of interest to compare this requirement with (4), which gives the production from the normal distribution of cosmic ray protons.

The total production of electrons with $\gamma > \gamma_1$, is

$$(21) \quad \sim \frac{2}{3} \cdot 10^{-22} n \gamma_1^{-\frac{1}{2}} \text{ cm}^{-3} \text{ s}^{-1}.$$

Inserting $n \simeq 10^{-2}$ atom cm $^{-3}$ for the halo, and equating to (20), gives $\gamma_1 \simeq 10^2 H^{-\frac{1}{2}}$. This is only weakly dependent on H , and for H of the general order of 10^{-6} G required $\gamma_1 \simeq 10^4$.

The significance of the previous paragraph is that electron production above $\gamma_1 \simeq 10^4$, arising from the normal cosmic ray protons, is adequate to give the correct emissivity, provided the magnetic intensity is high enough to produce synchrotron degeneration down to $\gamma \ll \gamma_1$, provided no electron secondaries are generated below γ_1 , and provided the necessary magnetic intensity turns

out to yield values of ν of the correct order, *i.e.* $\nu \simeq 10^8$ Hz. The synchrotron time scale $\sim 5 \cdot 10^8 H^{-2} \gamma^{-1}$ s. Eliminating γ with the aid of (7) enables this time scale to be written as $\sim 10^{12} \nu^{-\frac{1}{2}} H^{-\frac{3}{2}}$ s. If at $\nu \simeq 10^8$ Hz this is to be less than the age of the Galaxy, $\sim 4 \cdot 10^{17}$ s, and also less than the bremsstrahlung time scale, $\sim 3 \cdot 10^{15} n^{-1} \simeq 3 \cdot 10^{17}$ s for $n \simeq 10^{-2}$, we require $H > \sim 2 \cdot 10^6$ G.

The third of the three provisos mentioned above is easily checked, namely that the frequency $\nu \simeq 10^8$ Hz corresponds to a γ value below $\gamma_1 \simeq 10^4$. Inserting $\nu = 10^8$ Hz, $H > 2 \cdot 10^{-6}$ in (7) gives $\gamma < \sim 4000$. Still lower values of γ correspond to frequencies below 10^8 Hz.

The first of the three provisos is simply a demand that H be large enough. Setting $H \simeq 3 \cdot 10^{-6}$ G in the halo, instead of $3 \cdot 10^{-7}$ G, does not seem to be an embarrassment. In fact, the reverse is the case. A magnetic intensity as low $3 \cdot 10^{-7}$ G would not yield important coupling forces of an electromagnetic nature between the halo and the disk. From the point of view of our understanding of galactic structure this seems a retrogressive step. A field at least as high as $\sim 3 \cdot 10^{-6}$ G is to be preferred.

At first sight one may wonder at what essential point the present considerations differ from those of Section 1. The answer is in the proviso, still to be considered that no secondaries are produced below $\gamma_1 \simeq 10^4$. This allows the magnetic intensity to be increased sufficiently for synchrotron degradation to take place below γ_1 , leading to the spectrum $d\nu/\nu^{0.5}$. If the normal electron production rate (4) were to operate below γ_1 , we should simply obtain the spectrum $d\nu/\nu^{1.25}$ already rejected in (8). It is emphasized that the requirement is not for a mopping up of electrons below γ_1 (*e.g.* ionization losses), since such electrons are needed to give the radio spectrum—in fact the electrons must be controlled by the synchrotron process, not by any other process. The requirement is for an absence of production below γ_1 . It is difficult to see how this condition can be met without a drastic change in the cosmic ray picture developed in Section 1. This is why an entirely unambiguous determination of the radiofrequency spectrum is a matter of such great importance.

What at first sight appears to be a simple resolution of the issue is obtained by reducing the halo density, thereby reducing the production of secondaries over the whole range of γ . Electrons with $\gamma > \gamma_1$, could then be introduced as primaries. The difficulty with this suggestion is that cosmic rays returning to the disk from the halo would, during their sojourn in the halo, have passed through less than the 30 g of hydrogen calculated in Section 1. If n were reduced by a factor of 3 only the reduced mass of 10 g of hydrogen would be insufficient to destroy all (Li, Be, B) in the returning cosmic rays. A high ratio of (Li, Be, B) to (C, N, O) would be expected in these returning particles.

It could of course be argued that n is much lower still in the halo, $< 10^{-3}$ atom cm^{-3} , in which case the returning cosmic rays would have trav-

versed very little matter. But each time the cosmic rays return to the disk they traverse ~ 3 g of matter in the disk. There will be multiple returns, and cosmic rays will exist at any moment within the disk that have traversed unacceptable quantities of matter, ~ 12 g. If n is very small in the halo, we expect the maximum number of returns to be given by the ratio of the age of the Galaxy to TV_h/V_d . Now T has already been set at its upper limit, $\sim 7 \cdot 10^{13}$ s, and V_h/V_d can hardly be taken much greater than 10^3 giving TV_h/V_d not greater than $\sim 7 \cdot 10^{16}$ s. Within the age of the Galaxy we therefore expect at least 5 returns, giving cosmic rays that have traversed from ~ 3 to 15 g of hydrogen. This is inconsistent with the observed chemical composition.

To cope with this difficulty we could argue either, that during the characteristic time interval T for the disk, or during the time interval TV_h/V_d for the halo, there is an appreciable probability, 0.5 say, of escape into extragalactic space. Then the proportion of cosmic rays returning to the disk would fall off rapidly as the number of returns increased. But this at first sight simple resolution has now led to considerable complications, in which many of the attractive features of Section 1 have been lost. For example, the cosmic rays can no longer all start as heavy nuclei. Escape from the Galaxy is necessary. Primary electrons must be postulated.

The last of these complications can be avoided. And the artificial assumption that the probability of complete escape from the Galaxy just happens to be comparable with the probability of entering the halo (or of returning to the disk) can be avoided. The need in Section 1 to suppose that the probability of entering the halo was independent of rigidity, arose from accepting the heavy nuclei as the true primaries. Once we are forced to drop this attractive possibility, a rigidity-dependent probability can be postulated. We now regard the probability as increasing with rigidity in a marked degree. Then we take the low-energy cosmic rays as escaping from the Galaxy and high-energy cosmic rays as entering the halo—the route out of the Galaxy is easier at low energies, the route into the halo is easier at high energies, say at rigidities above 10^3 GeV. Then we can revert to the halo electrons as secondaries, since these now arise at $\gamma > 10^4$.

This picture would be consistent with a fairly regular field in the spiral arms that opened into extragalactic space. Instead of a diffusion problem, one has now to consider drift in magnetic inhomogeneities—even in a tolerably regular field there would still be some inhomogeneities. The drift velocity $\sim cR/L$, where R is the Larmor radius and L is the characteristic scale of the inhomogeneity. The time scale for drift through a distance L is therefore $\sim L^2/cR$. Setting $L \simeq 30$ parsec $\simeq 10^{20}$ cm, comparable with the half thickness of the disk, and $R \simeq 10^{15}$ cm for a cosmic ray of rigidity 10^3 GeV in a field $\sim 3 \cdot 10^{-6}$ G, this time $\sim 3 \cdot 10^{14}$ s, which is not much different from the above

value for T . Hence it is not out of the question to suppose that particles with rigidities of this order might succeed in escaping from a rather regular disk field into the halo.

The gains and losses of our new picture may be summarized:

Losses:

- 1) Cosmic rays cannot all originate as heavy nuclei.
- 2) There is no immediate explanation of why the energy of the proton component of the cosmic rays is comparable with that of nuclei with $Z \geq 6$ (unless this can be demonstrated to be a consequence of the chemical composition of the outer layers of type II supernovae).
- 3) In order to leave the Galaxy in $\sim 10^{14}$ s, particles must stream more or less smoothly out of the Galaxy. This could give a sense to the flow leading to spatial anisotropy. (Such anisotropy would hold only for rigidities $< 10^3$ GeV. At higher rigidities the particles enter the halo, and they return to the disk after a time interval $\sim TV_h/V_d$. On their return such particles can be expected to show a close approximation to isotropy. So anisotropy is not to be expected in air shower experiments. At low rigidities, isotropy might be introduced by magnetic fields within the solar system. The issue of anisotropy is not crucial, therefore, although it constitutes an awkward point.)

Gains:

- 1) The radio spectrum can be $d\nu/\nu^{0.5}$. (Naturally if this does not turn out to be necessary, the whole of the present section becomes redundant.)
- 2) The magnetic field in the halo can be taken of sufficient strength to give appreciable coupling forces between the disk and the halo.
- 3) The upper limit for T , set in Section 1 by the (Li, Be, B) to (C, N, O) ratio, is explained in terms of the time of travel out of the Galaxy. Indeed, the reason why most of the cosmic rays have not passed through more than ~ 3 g of hydrogen is very simple, because in the main cosmic rays leave the Galaxy before passing through more than this.
- 4) Hydrogen is detected spectroscopically in the expanding envelopes of type II supernovae. Hence it may really not be permissible to argue that protons are an unimportant component of the true primaries. This point goes some way towards neutralizing the first of the above «losses».
- 5) Photographs of extragalactic nebulae suggest that spiral structure is an ordered phenomenon. A reasonable degree of smooth streaming of cosmic rays along the arms is therefore to be expected, rather than a chaotic diffusion

picture. Reversals of the motions of individual particles at constrictions of the field can take place a moderate number of times without serious inconsistency arising between T and the time of travels, which $\sim 2\pi rc^{-1}$ multiplied by a factor ~ 5 for the effect of spiralling about the field lines and for the number of turns of the spiral arms from the sun to the edge of the Galaxy (about two such turns). Thus $T \sim 7 \cdot 10^{13}$ s and $10\pi rc^{-1} \simeq 3 \cdot 10^{13}$ s.

To finish this discussion, it may be noted that the above possibility leads to an electron flux that is appreciably less than the value calculated in Section 1. The flux is now mainly given by the secondaries generated during the passage of protons out of the Galaxy, *i.e.* only over the time scale T . Taking $n \simeq 1$ atom cm^{-3} for the disk, and multiplying (4) by $Tn \simeq 10^{14}$, gives $\sim 10^{-8} d\gamma/\gamma^{-\frac{1}{2}}$ for the electron distribution. Integrating from γ to ∞ , we have $\sim \frac{2}{3} \cdot 10^{-8} \gamma^{-\frac{1}{2}}$ for the density per cm^3 of electrons above. In terms of flux this corresponds to $\sim \frac{5}{3} \cdot 10^5 \gamma^{-\frac{1}{2}}$ or if the electron energy is expressed in GeV to $\sim 2E^{-\frac{1}{2}} \text{ m}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$ for the flux above energy E .

The effect is therefore to reduce the expected flux by a factor ~ 10 or more. If the flux should in fact turn out to be small, of order 0.1% of the proton flux at $E \simeq 1$ GeV, this would be a strong confirmation of the present rather unorthodox ideas. These ideas would also require a low value for the ratio of ^3He to ^4He .

On the present basis the distributed radio emissivity per unit volume from the disk must be low. The fact that the emissivity is apparently higher than the halo must be attributed to a considerable multiplicity of discrete sources, greatly expanded supernova remnants. The evidence on this point does indeed suggest that a considerable fraction of the radio emission of the disk arises from discrete non-thermal sources, probably from supernova remnants.

Extragalactic cosmic rays. — Supernovae of type II occurring in other galaxies may also be expected to produce cosmic rays. Since type II supernovae seem to be stars of large mass, they can be found only in galaxies possessing gas and dust, essentially in the spirals. It is therefore of interest that normal radio emission from galaxies is largely confined to Sb and Sc spirals, exactly as we should expect. Particularly, we should expect little in the way of radio emission from E-type galaxies, since type II supernovae do not seem to occur in them. Moreover, in comparing galaxies containing similar stellar populations, the supernovae activity depends simply on the numerical strength of the population in question. Since the optical emission also varies directly with the population, the supernovae activity—and hence the radio emission—will vary from galaxy to galaxy in the same way as the optical luminosity. This appears to be the case for Sb and Sc galaxies.

The close correspondence between the radio magnitudes and the optical magnitudes of Sb and Sc galaxies would indeed strongly suggest a stellar process

of origin of cosmic rays. It suggests, moreover, that the properties of the stellar assemblies in such galaxies are quite remarkably similar.

But this leaves the origin of the cosmic rays responsible for the abnormally strong radio sources quite unexplained. The requirement is for cosmic rays with total energy $\sim (10^{58} \div 10^{60})$ erg to be made available in a time interval no longer than $(10^7 \div 10^8)$ years. A supernova rate of at least 1 per year would be required, and this is not observed to occur in any galaxy. Moreover, the strong radio sources appear to occur mainly among *E*-type galaxies, and this is not at all what we would expect on the supernova theory. The implication is that the most intense processes of origin of cosmic rays are not at all of the kind described above. The optical evidence associated with the radio source M87 in the Virgo Clouds points to the same conclusion. This galaxy possesses a jet emerging from the nucleus, suggesting a large-scale activity affecting the whole galaxy, not a process arising from many individual stellar units. In this connection it may be remarked that M87 is not intrinsically a particularly strong source. The typical strong source is more intense by an order of magnitude at least.

A puzzling problem arises from the spectra of the strong sources, which are often about $d\nu/\nu^{0.8}$, very close to the spectrum calculated in Section 1, suggesting that the radio emission arises from cosmic rays possessing an energy spectrum similar to that found in the Galaxy. One wonders why two apparently quite distinct phenomena should produce nearly identical spectra. From this as a starting point it is possible to follow two quite distinct paths.

Referring back to the discussion of Section 1, a remarkable feature of the spectrum law was the elimination of most of the factors that related to the cosmic ray source itself. Indeed the only requirement was for comparable contributions from lifting the energies of already-existing particles and from the injection of new particles. This was the ratio K/k and the integral energy spectrum had index $1+k/K$. Any process that supplied both injection and lifting in a ratio close to unity would lead to a cosmic ray energy spectrum with index ~ 2 , and this would give reasonable agreement with the radio data.

A second possibility is to argue that the cosmic rays are not a galactic phenomenon at all, that the particles observed in the vicinity of the Earth have entered the Galaxy from outside, that there is just one universal distribution of cosmic rays possessing a unique energy spectrum. If it should turn out that the radio spectrum of the Galaxy really is $d\nu/\nu^{0.5}$, and if the non-solar electron flux should turn out to be low, this will become a more attractive possibility than it is at the moment. If it should be necessary to regard cosmic rays in the galactic disk as having a more or less free passage out into extragalactic space, it may well become more natural to regard the cosmic rays as having entered the Galaxy from outside, the time scale for entry and exit being $\sim T$. This would help in dealing with the anisotropy issue, and it would

remove a possible difficulty concerning cosmic rays with energies of $\sim 10^9$ GeV. This difficulty has not been mentioned so far, because of the present uncertainty as to the identity of these very-high-energy particles. If they should be protons, not heavy nuclei, and if the energy spectrum should turn out to extend to 10^{10} GeV, the difficulty of confinement inside the Galaxy becomes acute, particularly if the magnetic intensity has to be set as low as the halo value of $3 \cdot 10^{-7}$ G obtained in Section 1.

The usual reason for rejecting an extragalactic origin for the cosmic rays is the energy problem. How can one provide 1 eV cm^{-3} throughout the whole of space? This is clearly a cosmological question. In the steady-state cosmology there is no serious difficulty, at least in principle. The mean intergalactic mass density is $\sim 10^{-29} \text{ g cm}^{-3}$. If matter is created in this cosmology in the form of neutrons [5], the thermal energy density arising from neutron decay is $2.5 \cdot 10^{-12} \text{ erg cm}^{-3}$.

The velocity of sound, U say, in this hot gas sets a condensation distance scale, UH^{-1} , where H is the Hubble constant. Compression of cosmic rays during the formation of condensation lifts their energies. Each particle gains energy during compression according to an equation of the form of (15). The situation is similar to that described by Fermi in his second theory, where all collisions are of a « head-on » type. The condensation process causes cosmic rays to be reflected by an inward moving boundary, provided the cosmic rays are confined for an appreciable time within the condensation by a magnetic field. The gain rate α is of the order of the velocity of compression U , divided by the scale of the condensation, UH^{-1} . Hence α is just of order H . If we regard the expansion of the universe as providing the equivalent of Fermi's loss rate, then the constant β is also of order H , and the ratio $\beta/\alpha \sim 1$. Hence the required spectrum is obtained from rather cogent general considerations. The problem of injection remains, however, and is not so easily dealt with. The possibility of a universal injection process of an adequate strength turns on the existence of suitable plasma oscillations in the hot intergalactic gas.

3. — Cosmic rays of solar origin.

A crucial difference arises between cosmical and terrestrial electrodynamics, because characteristic decay times, given by $\sigma L^2/c^2$ (σ the conductivity, L the characteristic linear scale), are in general very long in the former and very short in the latter. Indeed, for most purposes in cosmical electrodynamics we can regard $\sigma L^2/c^2$ as so large as to be effectively infinite, and $\sigma L^2/c^2$ infinite implies that the ionized material and the lines of magnetic force are frozen together. Thus if the matter is moved by some non electromagnetic agency the lines of force will follow the motion. This permits the field lines to be distorted

and twisted into complex patterns that cannot be realised on the Earth. It also permits efficient storage of energy in the field, since in many cases the distortion and twisting leads to an increase in the integral of H^2 . An important feature of this process is that storage can take place in a volume V as a consequence of the motion of material outside V . Suppose a longitudinal field passes through two dense patches of gas, one of which is rotating about an axis parallel to the field. The magnetic lines of force are held by the non-rotating patch, while they are turned by the rotating patch. This leads to a twisting of the field in the region between the patches. Electric currents must flow in the region between the patches in order to represent this twisting in accordance with $\text{curl } \mathbf{H} = 4\pi\mathbf{j}/c$, but since the conductivity is essentially independent of density only very little material need exist in the intermediate region in order to provide for the current flow. The magnetic fields transmits a stress from one patch to the other, tending to transmit angular rotation. However, provided the patches are sufficiently dense, the magnetic stress will not equalize the rotations over a limited time scale. (Some such twisting effect almost certainly occurs in the solar atmosphere. Magnetic loops may extend high into the corona—in some cases loops may indeed emerge into interplanetary space. Such loops become twisted in the region of the corona and chromosphere due to the dense photospheric material rotating at the roots [6]). This is a particular case of the general proposition mentioned above, that the magnetic energy can be increased in a spatial region due to mass motion outside the region. It is possible to build comparatively large magnetic energies in regions where the mass density is small. Hence if some way can be found of discharging the accumulated magnetic energy, the small quantity of material in the region receives a large amount of energy per unit mass. This is precisely the requirement for the generation of high-speed particles. It is important to notice that any such discharge process cannot be one that is easily set up, otherwise the magnetic energy would never have been stored in the first place. For this reason the discharge process, if it occurs—and the case of the solar flare encourages us to suppose that it does—is likely to be a somewhat unusual one. It need occasion no particular surprise therefore that this aspect of cosmical electrodynamics is still in a very incomplete state and that speculations on the nature of discharges are still tentative.

The simplest attack on the discharge problem is to accept the decay time $\sim \sigma L^2/c^2$ and to argue that under rare conditions L can become very small. It is possible to find conditions [6] under which two distinct electromagnetic structures become aligned with the magnetic directions antiparallel. A surface must then exist between them on which $H = 0$. Such an arrangement is electromagnetically attractive (the opposite case, of magnetic directions parallel, is repulsive). Hence the structures press together. Mutual annihilation can only be prevented by a layer of gas between the structures in

which a gas pressure comparable to the magnetic pressure is developed. For mainly unionized hydrogen of particle density n within the layer this requires

$$(22) \quad H^2 \simeq 8\pi n k T,$$

where k is the Boltzmann constant and T is the temperature. Write $\bar{n}$ for the density of the uncompressed gas. Provided $\bar{n} \ll n$ a considerable degree of compression is necessary to produce the required layer density n . The layer must then be thin compared to the general dimension D , say, of the electromagnetic structures themselves.

To be more precise than we have been so far, the total rate of ohmic dissipation for a layer of width $2L$ is $\sim c^2 H^2 (8\pi^2 \sigma L)$, and the electric field component parallel to the layer, necessary to maintain the current, is $\sim cH/4\pi\sigma L$. Charged particles in the sheet $H=0$ are accelerated by this electric field. While such particles may well leave the sheet, curvature of their paths in the increasing magnetic field as they go away from the sheet always operates to bring them back to the sheet. For a field symmetrically distributed about $H=0$ the particles oscillate from one side of the sheet to the other, along a path that discharges continuously in the electric field. Hence if the particles are not appreciably deflected by collisions with other particles, and if their Larmor radii do not appreciably exceed L (in which case they spend most of the time in regions outside the layer where the electric field is much lower than it is inside the layer) there will be a continuous gain of energy up to a value of order

$$(23) \quad ecHD/4\pi\sigma L.$$

The condition that the accelerating particles are not appreciably deflected by collisions with other particles is a serious one that will be examined below. The condition on the Larmor radius limits the particle energies. Thus for protons we require

$$(24) \quad R < 3 \cdot 10^{-7} HL,$$

where R is the rigidity in GeV, H is in gauss, and L is in cm. We shall be concerned with cases where $H \simeq 10^3$ G, $L \simeq 10^3$ cm, in which case $R < \sim 300$ MeV. Strictly acceleration is not entirely absent at high values of R , the efficiency is reduced. Acceleration up to ~ 1 GeV, or even higher are not therefore prohibited, although (24) suggests that the majority of accelerated particles will have rigidities of a few hundred million volts, as indeed the solar cosmic rays mainly have. A further consistency is that insertion of $H \simeq 10^3$ G, $L \simeq 10^3$ cm, $D \simeq 10^9$ cm, $\sigma \simeq 10^{13} \text{ s}^{-1}$ in (23) gives an energy ~ 100 MeV, which corresponds closely to the rigidity ~ 300 MeV, given by (24).

The following discussion establishes that L can be of order 10^3 cm and

considers the condition under which collisions of an accelerating particle with other particles can be neglected. We consider first the value to be associated with L . We note that the gas in the layer will tend to be squeezed out of the layer by excess pressure. It will also tend to slide under gravity. The outflow from the layer takes place not over its two main faces which have areas $\sim D^2$, but over its rim with area $\sim 2LD$. Writing V for the outflow velocity, the rate of loss of material from the layer is $\sim 2nLDV$. A quasi-equilibrium is maintained, however, through material added at the main faces. This takes place at a rate $\sim \bar{n}D^2v$, where v is the electromagnetic drift velocity given by $v \simeq c^2/2\sigma\pi L$. Equating the rates of gain and loss gives

$$(25) \quad L^2 = \frac{\bar{n}c^2 D}{4\pi\sigma n V}.$$

The velocity V will in general be of order 10^6 cm s^{-1} in the chromosphere of the sun. This is of the order of the speed of sound, with which gas may be expected to stream under excess pressure. It is also of the order of the speed of a slide under gravity.

For $\bar{n}$ we may reasonably take $\sim 3 \cdot 10^{11} \text{ atom cm}^{-3}$ for the regions in which flares occurs. With $D \simeq 10^9 \text{ cm}$, $\sigma/c^2 \simeq 10^{-8} \text{ cm}^{-2} \text{ s}$, we have $L \simeq 10^{11} n^{-1/2}$. Inserting $H \simeq 10^3 \text{ G}$, $T \simeq 10^4 \text{ }^\circ\text{K}$ in (22) gives $n \simeq 3 \cdot 10^{16} \text{ atom cm}^{-3}$. Hence L is actually less than 10^3 cm . Indeed our formulae give $L \simeq 3 \cdot 10^2 \text{ cm}$. With so small a value, (23) gives acceleration to a rigidity of $\sim 1 \text{ GeV}$, which is comfortably near the general upper limit of energy of the majority of solar cosmic rays. It is true that (24) suggests that particles of this energy will tend to pursue orbits that take them somewhat outside the layer for most of the time. But this will not be true for all particles, although it may be true for most of them.

Other quantities are also of the required order of magnitude. So far, we have simply considered a general dimension D for the electromagnetic structures. In the flare case a reasonable structure might be a tube of radius $\sim 3 \cdot 10^8 \text{ cm}$ and length $\sim 10^{10} \text{ cm}$, with total volume $\sim 3 \cdot 10^{27} \text{ cm}^3$. The total magnetic energy for $H \simeq 10^3 \text{ G}$ is $\sim 10^{32} \text{ erg}$. The time scale of dissipation, for two tubes pressed together (with common generators) is $3 \cdot 10^8/v$, where v is the electromagnetic drift velocity given above as $c^2/2\pi\sigma L \simeq 10^5 \text{ cm s}^{-1}$. Thus the dissipation time scale $\sim 3 \cdot 10^3 \text{ s}$ — about 1 hour, which is a reasonable estimate for a strong flare.

Before proceeding to the collision problem it is necessary to consider an issue that might vitiate the above calculation. If T were to rise appreciably above $10^4 \text{ }^\circ\text{K}$, it would no longer be correct to set $\sigma \simeq 10^{13} \text{ s}^{-1}$, since $\sigma \propto T^{3/2}$ and a value $\sigma = 10^{13} \text{ s}^{-1}$ refers to T somewhat below $10^4 \text{ }^\circ\text{K}$. Writing x for the fraction of the hydrogen within the layer that is ionized we note that the

rate of radiation by the hydrogen in the layer $\sim 10^{-24} n^2 x^2$ erg cm $^{-3}$ s $^{-1}$ in the Balmer continuum. We take the layer to be optically thin to this radiation—for if it were not the radiation rate would approach the thermodynamic value, which would be greater still. Thus the radiation rate per unit area of the layer $\sim 2 \cdot 10^{-24} n^2 x^2 L$. If this rate is adequate to radiate away the heat released per unit time by the ohmic losses, which $\sim c^2 H^2 / 8\pi^2 \sigma L$ per unit area, without x needing to become of order unity, then we can regard T as not exceeding 10^4 °K (since x rises to order unity as T increases above 10^4 °K). Equating the radiation rate to the ohmic loss rate gives

$$(26) \quad nx \simeq 10^{11} \cdot \frac{H}{L} \left(\frac{c^2}{\sigma} \right)^{\frac{1}{2}} \text{ ionized atom cm}^{-3}.$$

With $H \simeq 10^3$ G, $L \simeq 3 \cdot 10^2$ cm, $(c^2/\sigma)^{\frac{1}{2}} \simeq 10^4$ cm s $^{-\frac{1}{2}}$, $n \simeq 3 \cdot 10^{16}$ atom cm $^{-3}$, (26) gives $x \simeq 0.1$, and the required condition is satisfied by a safe margin.

It is also satisfactory that $c^2 H^2 / 8\pi^2 \sigma L \simeq 3 \cdot 10^9$ erg cm $^{-2}$ s $^{-1}$ for the radiation rate per unit area of the layer. This again is a reasonable value for a strong flare.

Turning now to the collision problem, the mean-free-path for protons is determined essentially by collisions with other protons. The mean-free-path for electrons is less than that for protons since electrons are deflected both in collisions with other electrons and also in collisions with protons, whereas electron-proton collisions do not produce much deflection of the protons. For this reason, electrons are less likely to be accelerated than protons.

A proton of energy η has a free path $\eta^2 / (2\pi e^4 n x \log A)$, where $\log A$ is a slowly varying quantity that takes a value ~ 5 in circumstances relevant to solar flares. The condition for a proton to discharge continuously in an electric field $\mathcal{E}$ is

$$(27) \quad e\mathcal{E}\eta^2 / (2\pi e^4 n x \log A) \geq \sim \eta.$$

Considering the limiting case given by the equality sign, inserting (26) for the product nx , and setting $\mathcal{E} \simeq cH/4\pi\sigma L$, gives

$$(28) \quad \eta \simeq 5 \cdot 10^{12} e^{\frac{1}{2}} \sigma^{\frac{1}{2}} \log A.$$

With $\log A = 5$, we hence have $\eta \simeq 5$ kV. The conclusion therefore is that protons with energy above 5 kV discharge continuously, attaining eventual energies upward of 100 MeV. Protons with energies below 5 kV do not discharge, being prevented from doing so by collisions with other protons.

The final question remaining for consideration is how protons with initial energies in excess of 5 kV can arise in the layer. Although only a very small

fraction of the particles of the layer need be accelerated to high energy, this small fraction cannot immediately be supplied by the high-energy tail of the Maxwell distribution. It can be shown that not more than about 1 proton in $\sim 10^{10}$ needs to be in the process of acceleration at any particular moment—otherwise the electric current arising from the accelerated particles would alone become greater than $cH/4\pi$, in which case a back electric field would reduce value of $\mathcal{E}$ below $cH/4\pi\sigma L$ —whereas $\exp[-\eta/kT]$ for $\eta \simeq 5$ kV, $kT \simeq 1$ V, is clearly very much less than this.

At first sight the injection of protons might therefore appear to raise a difficult problem. It is unrealistic, however, to suppose that the layer remains everywhere of the same thickness. Irregularities may be expected to develop, L being larger in some places than in others. Again, at first sight, this does not seem of help in the problem, since (28) is independent of L . The result (28) was derived from (27) by using the value (26) for nx , however, and the value (26) for nx depended upon a steady-state in which the radiation rate equalled the ohmic dissipation rate. If L were to decrease very quickly at a particular place on the layer, so quickly that a lag in the ionization level occurred, then $\mathcal{E}$ would increase in (27) while x would remain constant. The limiting value of η would then be decreased proportionately to L .

It remains to show that such a lag in the ionisation rate can occur. The total mass of hydrogen per unit area of the layer is of order $10^{-24}nL$ g cm $^{-2}$. The amount of energy required to ionise a fraction x of this hydrogen is $10^{-11}nxL$ erg cm $^{-2}$. The ohmic dissipation rate $c^2H^2/8\pi^2\sigma L$ per unit area. If the whole of the energy released by ohmic dissipation were taken up in ionizing hydrogen, the time required to change x appreciably would be $\sim 10^{-9} \cdot nx\sigma L^2/c^2H^2$. With nx given by (26) this is $\sim 10^2(\sigma/c^2)^{1/2}L/H = 10^{-2}L/H$, which can be compared with the time required for irregularities to develop in the layer. The latter time scale we expect to be of order L/V , V being the streaming velocity in the layer. With $H \simeq 10^3$ G, $V \simeq 10^6$ cm s $^{-1}$, the ionization time scale is ~ 10 times longer than the time scale for development of the irregularities. Since both time scales decrease linearly with L it appears that the layer thickness could decrease locally and momentarily to a very small value, in which case particles would be accelerated locally above the required threshold. Such particles continue to be accelerated on entering regions of the layer where L takes its normal value.

It is emphasized that L cannot remain for very long at a small value. Eventually x must rise to a value of order unity. This implies a rise of T , leading to an increase of σ , and hence an increase of L in accordance with (25). For this reason L can only decrease momentarily to a value close to zero.

Given sufficient injection of particles above the threshold, the current carried by the accelerated particles rises to a value $\sim cH/4\pi$. At this stage a back electric field prevents the acceleration from continuing at its former

rate. This is achieved through a reduction of $\mathcal{E}$ and through a consequent increase in the acceleration threshold. The situation is that magnetic energy is then converted with high efficiency into cosmic rays. The reduction of $\mathcal{E}$ causes the normal ohmic heating to decrease, so that the emission of radiation from the layer is decreased. Cosmic rays are produced at the expense of radiation, which falls below the value $\sim 3 \cdot 10^9 \text{ erg cm}^{-2} \text{cs}^{-1}$ estimated above.

It is remarkable that the determination of the acceleration threshold from (28) does not involve H , L , D , or $\bar{n}$. Apart from small changes in the factor $\log A$, the acceleration condition would be exactly the same for a magnetic reservoir of galactic dimensions provided T remained of order 10^4 °K. This strongly encourages the belief that flare discharges might occur over much greater dimensions than those of the Sun.

REFERENCES

- [1] M. V. K. APPA RAO: Preprint (1961).
- [2] C. L. CRITCHFIELD E. P. NEY and S. OLEKSA: *Phys. Rev.*, **85**, 461 (1952).
- [3] W. A. FOWLER and F. HOYLE: *Ap. Journ.*, **132**, 565 (1960).
- [4] F. HOYLE: *M.N.R.A.S.*, **120**, 338 (1960).
- [5] F. HOYLE: *11-th Solvay Conference* (1958), p. 66.
- [6] T. GOLD and F. HOYLE: *M.N.R.A.S.*, **120**, 89 (1960).

The U. S. Program for Particles and Fields Measurements in Space.

G. H. LUDWIG

National Aeronautics and Space Administration

1. - Introduction.

During the past several years, the United States has been heavily involved in a program of investigation of space phenomena by the use of earth satellites and deep space probes. The success of this program thus far is indicated by the number and the significance of the discoveries made during this period. This space science program is to a considerable extent an outgrowth of the program for investigating the upper atmosphere by the use of rockets and balloon-borne instrumentation. The possibility of the use of rockets for upper air research was recognized as early as 1919, by Dr. ROBERT H. GODDARD when he published his paper entitled *Method of Reaching Extreme Altitude* [1]. Plans were made early in 1932 for the use of Goddard's rockets during the Second Polar Year to lift small payloads of instruments to heights of up to fifty miles. But these hopeful schemes were interrupted by the economic depression. It was not until soon after the Second World War that rockets were actually used for such research. In October 1945, a small liquid-fueled research rocket known as the WAC-Corporal was fired to an altitude of 65.5 km with a 10 kg payload. Soon after this, on 16 April 1946, the first captured German V-2 was launched at the White Sands Proving Ground with one of Dr. JAMES A. VAN ALLEN's Geiger-Müller counters in the nose section to measure the intensity of Cosmic rays above the dense atmosphere. By January 1958, 68 V-2's, 63 Aerobees, and 7 Viking rockets had been fired into the high atmosphere. The rocket program continued to accelerate rapidly, and plans were made to include a large number of rocket launchings in the U.S. International Geophysical Year program. Concurrent with the development of these rockets and rocket techniques was the development of the miniaturized and highly rugged instrumentation needed for the performance of useful experiments. This

instrumentation was developed to the point where, in 1954, moderately complex Geiger-Müller counter arrays and scintillation detector systems were being carried in rockets having payload capabilities of only 5 to 10 kg. This work led directly to the development of the first highly complex transistor instrumentation which was developed during the years 1956-1957 for use in the early U.S. satellites.

Following the successful launchings of the first U.S. spacecraft—Explorers I, III and IV, and Vanguard I—the need for an integrated national program for conducting space research was obvious, and the National Aeronautics and Space Administration (NASA) was established to formulate and carry out this task. A recent revision of the original act defines the mission of NASA in more detail.

« The administration, in order to carry out the purpose of this act shall

1) formulate specific national objectives in space and, in the light of such objectives, develop a comprehensive program for the exploration, investigation, and utilization of space for peaceful purposes;

2) conduct research into problems of flight within and outside the earth's atmosphere with the view to their practical solution, including research in the field of aeronautics necessary to the continued advancement of both civilian and military aviation;

3) conduct such activities as may be required for the exploration, scientific investigation, and utilization of space for peaceful purposes, and develop space vehicles for use in such activities;

4) arrange for participation by the scientific community in planning scientific measurements and observations to be made through use of aeronautical and space vehicles, and conduct or arrange for the conduct of such measurements and observations and;

5) provide for the widest practicable and appropriate dissipation of information concerning the activities and the results thereof. »

Based on these general policies and objectives, the specific objectives of the NASA program are threefold (*):

1) To produce scientific data on the space environment, the sun, earth and planet, and the galaxy, using unmanned spacecraft equipped with instrumentation and telemetry to relay data to the ground. This information is essential to all utilization of space and to an understanding of the physical universe and its relation to man.

(*) See the proceedings of the NASA-Industry Program Plans Conference, July 28-29 (1960).

2) To study early applications of earth satellites to meteorological research and weather forecasting, long-distance wideband radio communication, navigation, and similar tasks.

3) Explore the problems connected with the travel of man in space, at first in orbital flight around the earth for short periods, later in flights to the moon, and still later to planets and the outer reaches of the solar system. The pace of manned exploration will be determined largely by the results obtained in the early orbital flights.

2. - Accomplishments of the first years.

With these objectives in mind, it is interesting to review the accomplishments of the first few years of the U.S. program. By 23 May 1961, nine Explorer satellites, four Pioneer planetary probes, and three Vanguard satellites had been successfully launched as a part of the U.S. space science program. Two Tiros satellites, and an Echo satellite have been launched as a part of the earth satellite applications program to fulfil the objectives listed in 2) immediately above. And the Freedom 7 Project Mercury capsule has been launched on a ballistic trajectory as one of the first steps in the program outlined in 3) above. These satellites and deep space probes have already provided a large amount of information about the space environment. Of the satellites launched thus far for the purpose of conducting space science experiments, many of them have been designed primarily for investigating energetic particles and magnetic fields in space. Of those instrumented for other purposes, Vanguard I was intended primarily to study the shape of the earth, and Vanguard II was instrumented for the investigation of the cloud cover over the surface of the earth. Explorers VIII and IX were developed to investigate the ionosphere, the density of the upper atmosphere, and passive communication techniques. Although the results obtained from the applications and manned satellite programs and the satellites carrying other than energetic particle and magnetic field experiments are interesting, they will not be discussed further in this paper.

Explorers I and III were launched on 1 February 1958 and 26 March 1958 respectively. Each of these satellites contained a single Geiger-Müller counter having a geometric factor of approximately 17 cm^2 . This counter can be seen in the mid-section of the photograph of Explorer I shown in Fig. 1. In this satellite, the Geiger-Müller tube counting rate was telemetered directly to the ground receiving stations. In Explorer III, in addition to being telemetered directly, the counting rate was also stored by a small tape recorder in the satellite and delivered to the ground receiving stations on command at the completion of each orbit. These two satellites discovered the existence of the

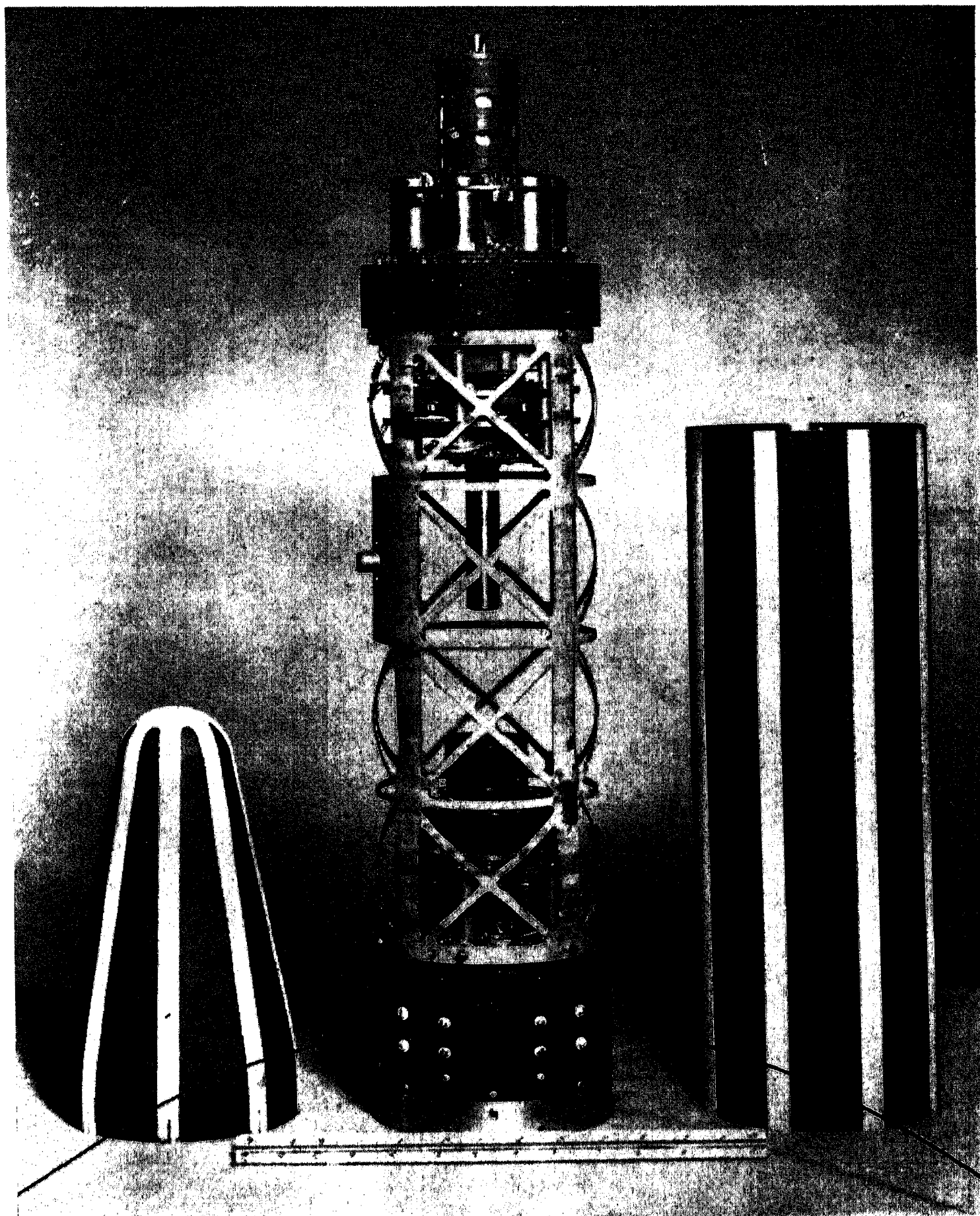


Fig. 1. - The Explorer I instrumentation section.

very high intensity geomagnetically trapped radiation now known as the Van Allen radiation. They provided the first rather crude survey of the spatial extent in this region of radiation [2]. Immediately following the discovery of this region of intense corpuscular flux, another satellite was prepared specifically for the purpose of further defining its characteristics. This satellite, known as Explorer IV, was launched on 26 July 1958. It provided: a more complete survey of the spatial distribution of the radiation at radial distances up to 2220 km and between latitudes of $\pm 51^\circ$ geographic; a measure of the radiation intensity by detectors having several different characteristics and under various amounts of absorber, thus providing several points on the particle integral energy spectrum curve at a large number of positions; a rough indication of the angular distribution of the trapped particles; and important information about the trapping mechanism obtained as a result of the artificial electron shells created by project ARGUS [3-7].

The experiments described above were carried in satellites having apogee heights of less than 2500 km. As a result, the configuration of the more distant regions of the belt structure was not known. The existence of two separate belts had not been unambiguously indicated by the low altitude data, although its possibility was realized. The importance of extending the survey to a greater radial distance was obvious. As a result, radiation detectors were prepared for Pioneers I, II, III and IV. The Pioneer II launch attempt was unsuccessful. Pioneers I and III traveled to distances of about 100 000 km, and the data obtained from these two spacecraft during their traversals of the radiation belt structure gave us our first information about the radial extent [8-10]. Pioneer I, launched on 11 October 1958, indicated that the high intensity radiation was confined to the region near the earth [11]. It did not, however, show the existence of the two distinct belts owing to gaps in telemetry recovery. Pioneer III, launched on 6 December 1958, made two cuts through the trapping region and clearly revealed the existence of the two great belts. Pioneer IV, launched on 3 March 1959 into a solar orbit, further confirmed the findings of Pioneer III and, by comparison of the data of Pioneer IV with that of Pioneer III, showed that the intensity of the outer belt was variable and apparently related to solar activity, thus indicating the high probability of the solar origin of the particles responsible for the outer belt. The two G.M. tubes can be seen protruding from the top of the Pioneer IV payload shown in Fig. 2.

Explorer VI, launched 7 August 1959, contained a proportional counter array, a Geiger-Müller counter, an ionization chamber, a scintillation detector, and a simple one-axis magnetometer. The data received from this satellite indicated the existence of a torroidal current ring located in the region from 5 to 7 earth radii. They also indicated that the magnetic field in space experiences widespread disturbances during the times of at least some mag-

netic storms, and that changes in the particle pitch angles (angles between direction of motion and the magnetic field lines) are followed by changes in

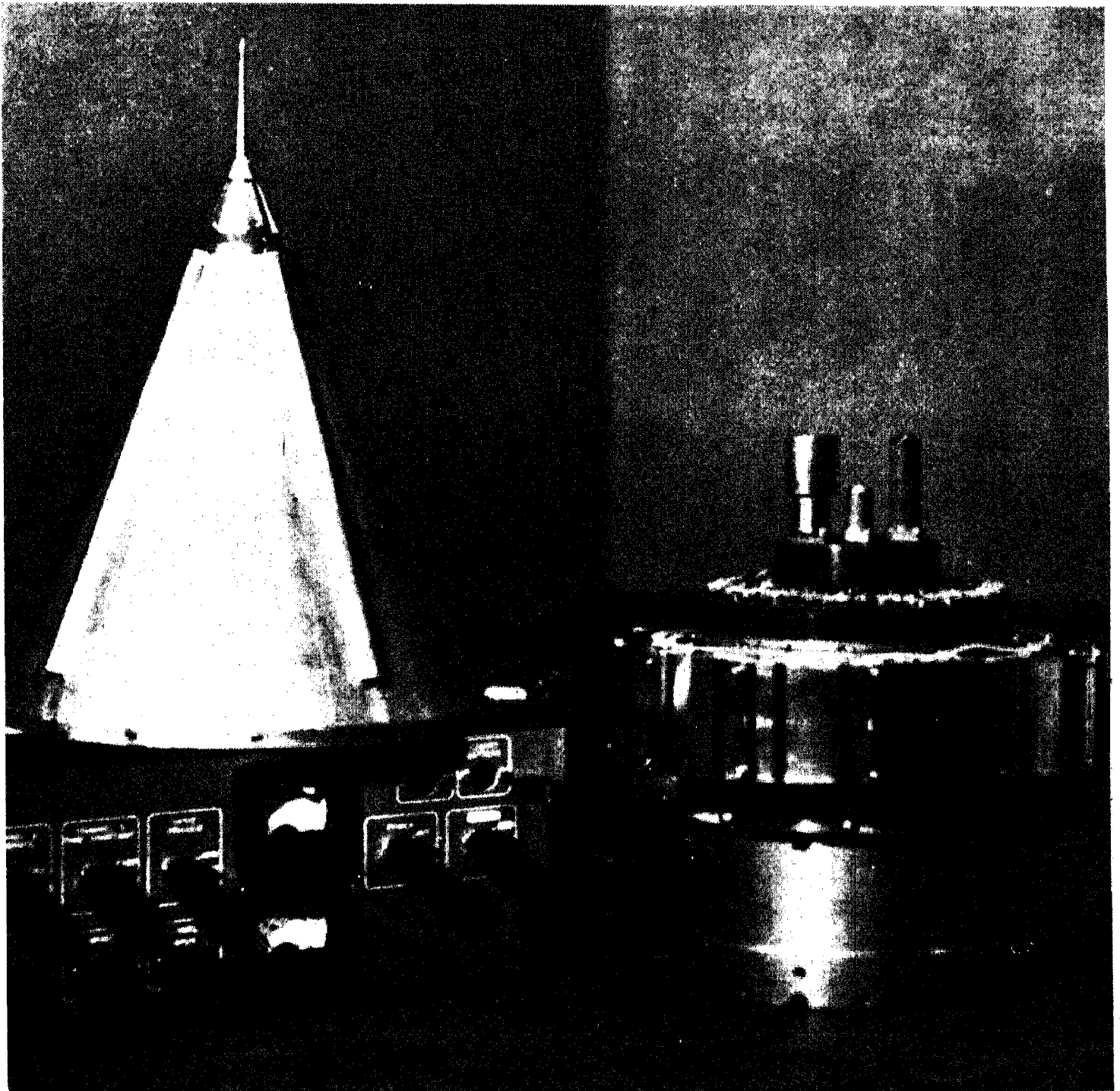


Fig. 2. - Pioneer IV. The two G.M. counters are located in the two outside vertical cylinders located on the top of the payload.

radiation intensity as observed by a given detector. In addition, the data showed the radial motion of the peak intensity in the outer belt at these times [12-17].

Explorer VII, (see Fig. 3) launched 13 October 1959 into a nearly circular orbit around the earth, and Pioneer V (see Fig. 4), launched 11 March 1960,

into a solar orbit, operated concurrently so that information was obtained at large radial distances at the same time that the intensity at the lower edges

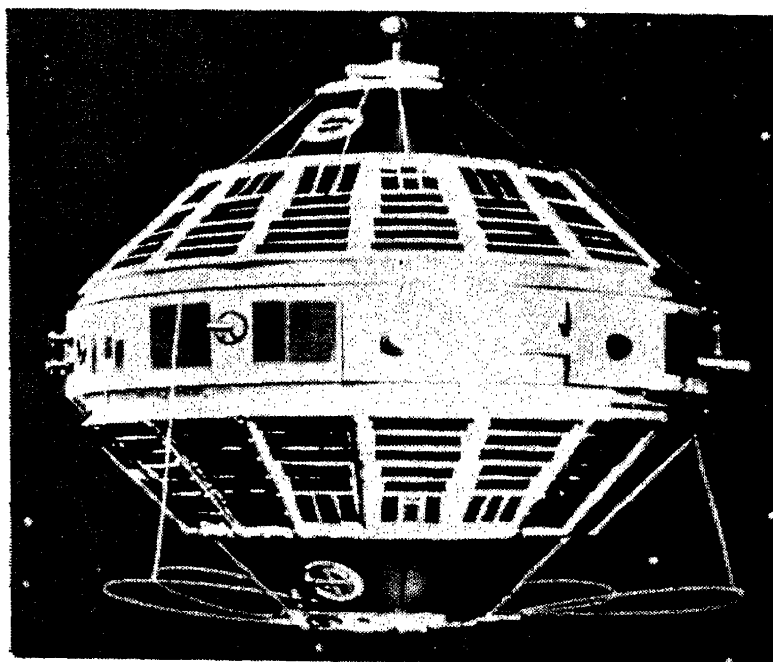


Fig. 3. - Explorer VII.

of the radiation belt structure was being monitored. Explorer VII carried two G.M. counters, one relatively unshielded and the other shielded by about 1.5 g cm^{-2} of material. Pioneer V carried a proportional counter array, a G.M. counter, an ionization chamber, and a search coil magnetometer similar to the Explorer VI instruments. Data from these two spacecraft provided a great amount of new information [18-29]. They indicated that the Forbush decrease is widespread, showing approximately the same decrease at $5 \cdot 10^6 \text{ km}$ from the earth as near the earth's surface. Solar

protons were detected on both Pioneer V and Explorer VII following the flare of 1 April 1960. The correlation of the Pioneer V and Explorer VII

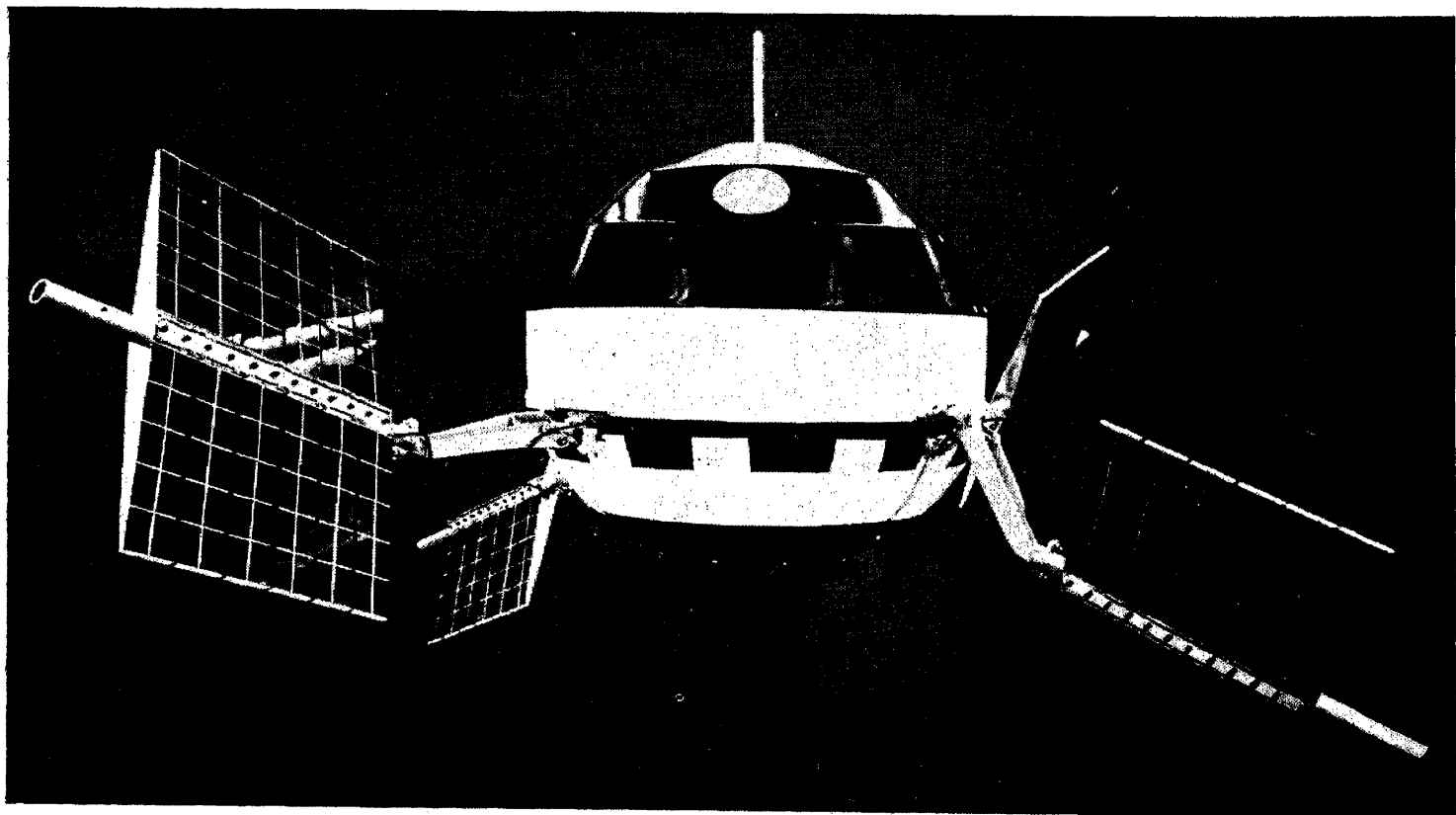


Fig. 4. - Pioneer V.

data indicate that the energetic electrons having energies greater than 30 keV which are observed in the outer belt following that solar disturbance must have been locally accelerated, since a sufficiently large integrated particle flux having that energy range was not detected passing the position of Pioneer V preceeding the enhancement of the outer belt intensity. Pioneer V, in addition, measured the intergalactic magnetic field. The value varied from 2.7γ during quiet periods to a maximum of 40γ during magnetically disturbed periods (one γ equals 10^{-5} G). The counters in Explorer VII operated until March 1961, providing a large amount of information about the adiation intensity in the lower fringes of the belt structure. They have indicated the high stability of the inner belt and have revealed many interesting spatial and temporal structural features of the outer belt. The ground observation of a red auroral arc ($6300 \text{ \AA}$) immediately below the region of the outer belt peak intensity as determined by Explorer VII following the solar event on 28 November 1959 indicated that particles being dumped from the outer belt are probably responsible for at least a portion of the auroral activity.

The Vanguard III satellite launched on 18 September 1959 carried a proton precession magnetometer (shown on the top of the boom in Fig. 5) which measured the magnitude of the earth's magnetic field throughout the orbit, which extended from 510 to 3700 km in range, and from north 33° to south 33° in latitude. The data obtained from this flight are still being analysed and are expected to result in an accurate

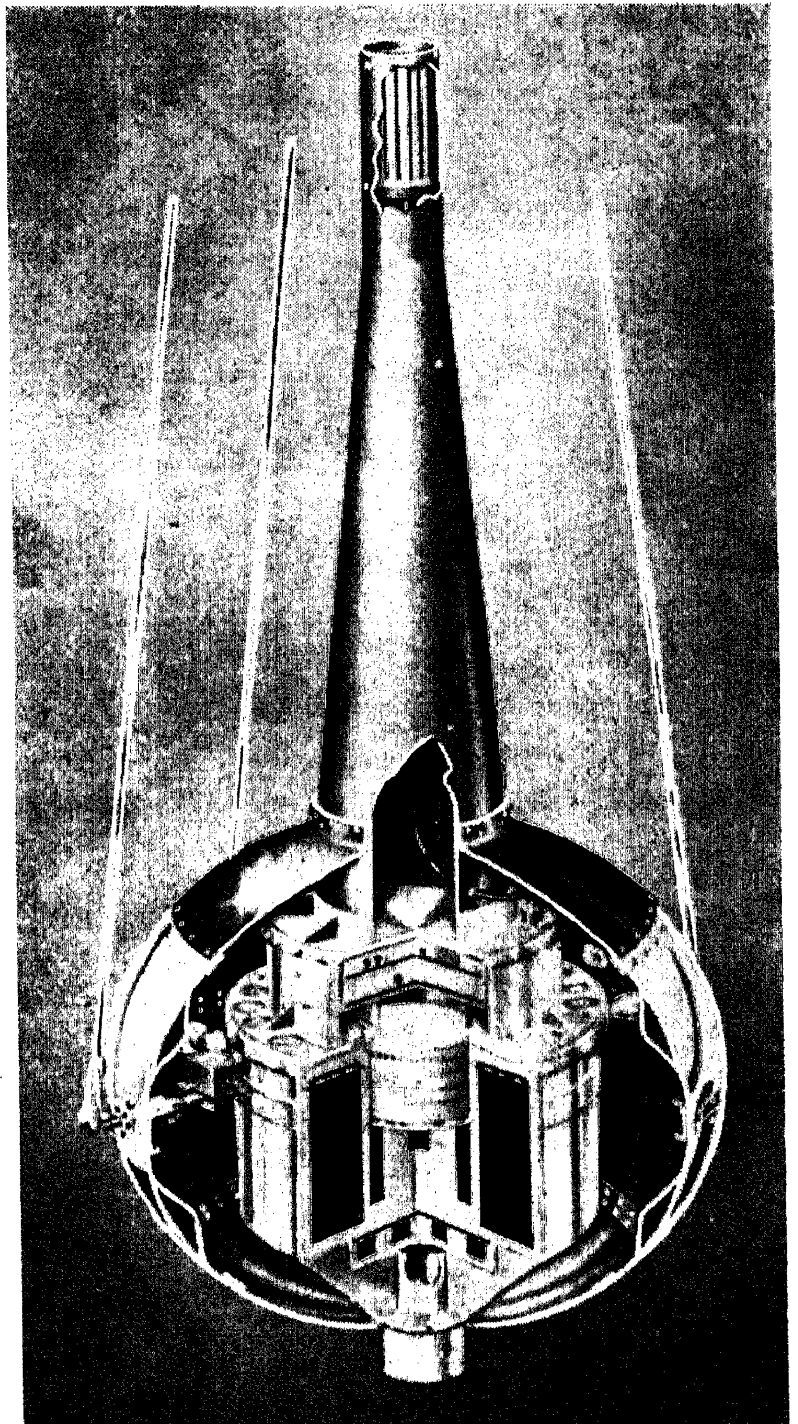


Fig. 5. - Cutaway view of Vanguard III.

mapping of the earth's magnetic field and in the production of a set of 1959 values for the 48 coefficients in the six term expansion of the earth's magnetic field. The data indicate that the earth's magnetic field at small radial distances is highly stable.

The next addition to the family of Explorer satellites which contained magnetic field and energetic particle experiments was Explorer X, launched on 25 March 1961. This spacecraft, shown in Fig. 6, carried an alkali vapor

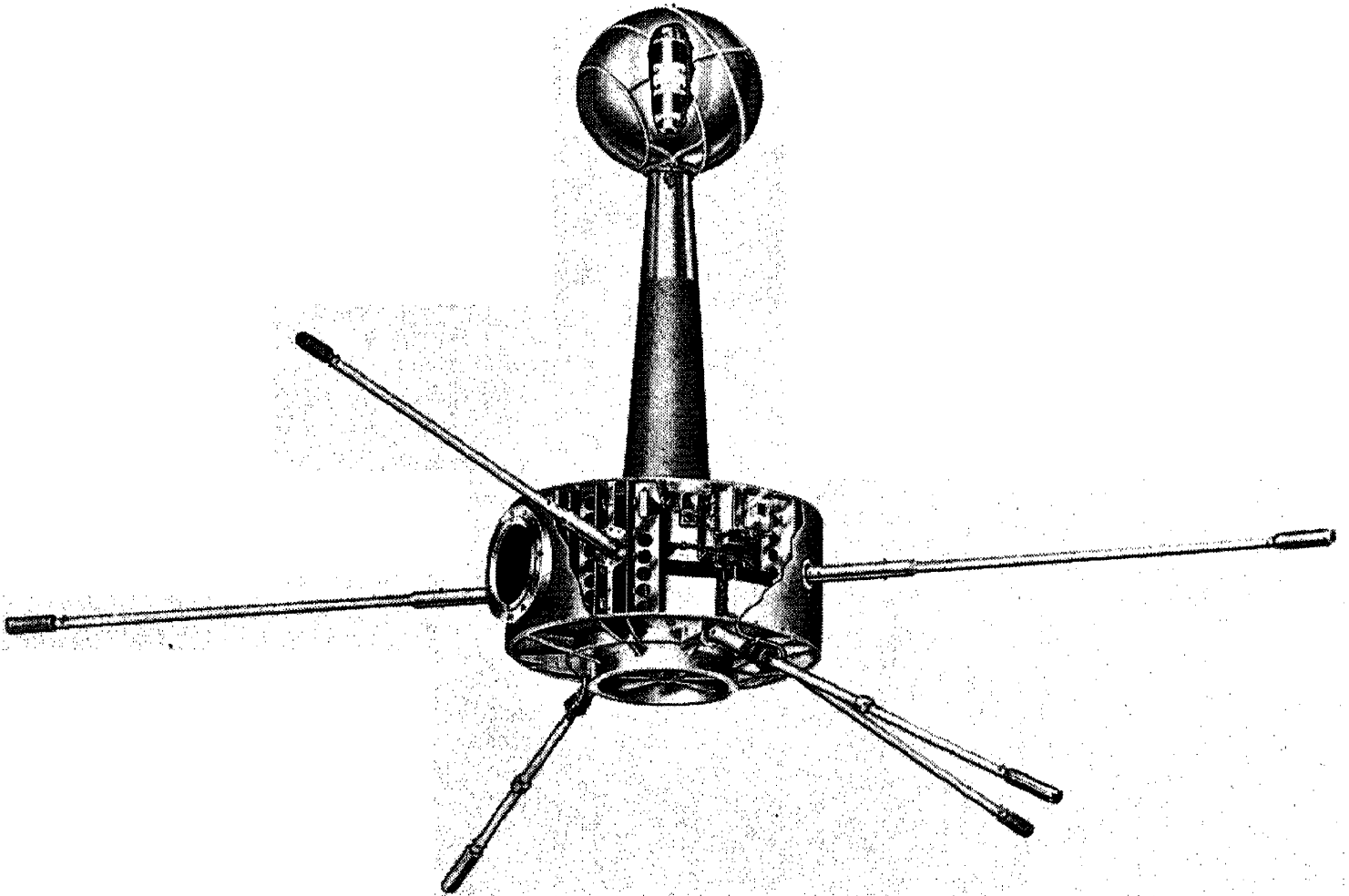


Fig. 6. — Cutaway view of Explorer X.

maser magnetometer, two flux-gate magnetometers, and a proton plasma probe covering the energy range from below 5 to 2300 eV, to a radial distance of about 260 000 km. These instruments indicated that the magnitude of the interplanetary magnetic field varied between 6 and 40 γ . Very little plasma was detected during the periods when the magnetic field was high at times prior to the sudden commencement which occurred near apogee. More information about both the magnetic field and the plasma is expected following the more complete analysis of the data.

Explorer XI was launched on 27 April 1961 to study the γ -rays. It carried

a complex γ -ray telescope employing scintillation and Čerenkov detectors in an array designed to discriminate against charged particles and neutrons. The data are now being reduced, and the results of the experiment should be available in several months.

In addition to the satellites and space probes discussed above, a number of sounding rockets have been fired. Some of the most recent firings include a series of Solar Beam Experiments flown from Fort Churchill during the September and November 1960 solar events, and a rocket flown into the lower edge of the inner belt from the Pacific Missile Range in California. These recoverable vehicles carried nuclear emulsions for the study of the Z and energy spectra and the fluxes of the electrons and protons. They have indicated that the proton energy spectrum on the inner edge of the inner belt has a slope which is dependent on the latitude [30], and have measured the energy and Z spectra of solar protons arriving at Fort Churchill following the solar flares of 3 September-12 November 1960 [31].

3. - Future energetic particle and magnetic-field investigations.

The term « Energetic Charged Particles » can be divided into four major groupings. The primary cosmic rays arise from acceleration mechanisms located outside the solar system but are modulated both in spectrum and intensity by solar system electrodynamics. Their study in the past has been primarily by means of high altitude rockets and balloons and on the earth's surface where atmospheric effects and variations in the earth's magnetic field due to ring currents and solar plasma have caused additional modulations. The use of satellites and space probes for the investigation of these particles permits the elimination of these local modulations so that the characteristics of the primary particles and of the solar system-induced variations may be studied more quantitatively.

The second major grouping consists of the energetic solar particles which are accelerated in or near the sun on the occasion of major solar flares. The relative abundance of the elements in these streams of particles is approximately the same as the relative abundance of elements on the sun. In addition to the nucleons, electrons and photons are also released during these flares. These streams of charged particles carry along their own magnetic fields which, upon approaching the earth, cause modulations of the particle fluxes reaching the earth from all external sources and of the structures of the trapped radiation belts.

The third major category is the continuously emitted low energy solar plasma, referred to as the « solar wind ». This plasma consists of particles having both positive and negative charges. It produces a charged particle

density in the region of the earth which is appreciably greater than that in interstellar space, and, due to its general motion outward from the sun, produces some distortion of the distant geomagnetic field and a resultant modulation of particles reaching the earth from outside the earth system.

The fourth major grouping is the reservoir of charged particles existing in geomagnetically trapped orbits near the earth and near at least some of the other planets. It appears at this time most probable that the electrons and protons existing within the earth's inner radiation belt result primarily from the decay of neutrons produced by cosmic ray interactions in the high atmosphere. Most of the particles in the outer belt most likely have their origin at the sun although it is probable that some local acceleration mechanism acts to inject the particles into their trapped paths and to produce the energy spectra observed.

From the above very brief introductory remarks, it is clear that a whole major field of investigation has explosively expanded during the last three years as a result of the availability of vehicles capable of carrying detectors into space. This field is actually a mixture of physics and astrophysics, and is directed toward the investigation of the basic questions of particle acceleration, the generation of magnetic fields in the stars, planets, and in galactic space, and the study of the motions of matter and magnetic fields in space. It is expected that the ultimate outcome of this line of investigation will be a better understanding of a large number of broad cosmological questions such as the development and dynamics of galaxies, the distribution of matter in the galaxies, the origin of the high energy cosmic rays, and the physics of stars and planets.

In order to gain this broad understanding, at least some of the following more specific investigations will need to be conducted by experiments in satellites and space probes in the future.

A) *Galactic cosmic rays.* In order to gain an understanding of the origin of cosmic rays and the acceleration mechanisms acting on them, it will be necessary to study:

- 1) the cosmic ray charge spectrum including, perhaps, the relative abundances of the various isotopes,
- 2) the energy spectra of the constituents,
- 3) the presence of and the energy spectra of high energy electrons, anti-matter, and other particles, and photons,
- 4) cosmic ray variations and the relationships between these variations and other events in the solar system,
- 5) the directional characteristics of the cosmic radiation.

It is also expected that the cosmic ray flux will continue to be used as a source of high energy particles for the study of high energy nuclear interactions and nuclear forces, and in the search for new particles.

B) Energetic solar flare particles. In order to study the production of the energetic solar particles and the magnetic fields which are intimately associated with their production near the sun, and to study the effects of these solar particles on the geomagnetic field and the earth's atmosphere, it will be necessary to continue investigation of:

- 1) the relationships between the relative abundances of the various elements in the solar particle stream and in the sun itself,
- 2) the particle fluxes as a function of time,
- 3) the energy spectra of the various constituents, and
- 4) correlations between the production of energetic solar particles and other phenomena such as visible flares, coronal streamers, solar radio noise, arrival of energetic solar particles near the earth, geomagnetic disturbances, variations in the structure of the radiation belts, and aurorae near the earth.

It should be noted, however, that it will not be possible to press vigorously investigation of these particles during the next three to six years, since we are approaching sunspot minimum and these events can be expected to be rare.

C) Solar plasma. To give insight into the physical processes occurring in the sun, to study certain modulations of the primary cosmic ray flux, and to determine the nature of some features of the geomagnetic activity, it will be necessary to investigate the low energy plasma emanating from the sun. The following investigations will be necessary:

- 1) determination of the fluxes and energy spectra of all components of the solar plasma,
- 2) study of the variations of the solar plasma at various positions in the solar system and at various times,
- 3) correlations with other solar and geophysical phenomena.

D) The trapped radiation. In order to gain a complete understanding of the source and loss mechanism acting on particles in the trapping regions, it will be necessary to gain a much more comprehensive understanding of:

- 1) the energy spectra of the various constituents of the inner radiation belt as a function of position,
- 2) the temporal characteristics of the inner belt,

3) the energy spectra of the constituents of the outer belt, with especial emphasis on continuation of the investigations to much lower energies,

4) the angular distributions of the various constituents of the outer belt as a function of position along the trapping lines,

5) the morphology of events following solar flares and their relationships to other solar activity and geophysical phenomena.

4. - The launching vehicles.

In order to understand more completely the capabilities and shortcomings of the spacecraft now being used and planned for the immediate future, it is necessary to review briefly the capabilities of the vehicles which will be used to place them in their orbits. In the vehicle program, and in the spacecraft program too, as will be evident later, several guiding principles apply:

1) The NASA space science program is a civilian program and, therefore, may differ somewhat in approach from the one that is based on military requirements. It is NASA's responsibility to obtain the greatest possible

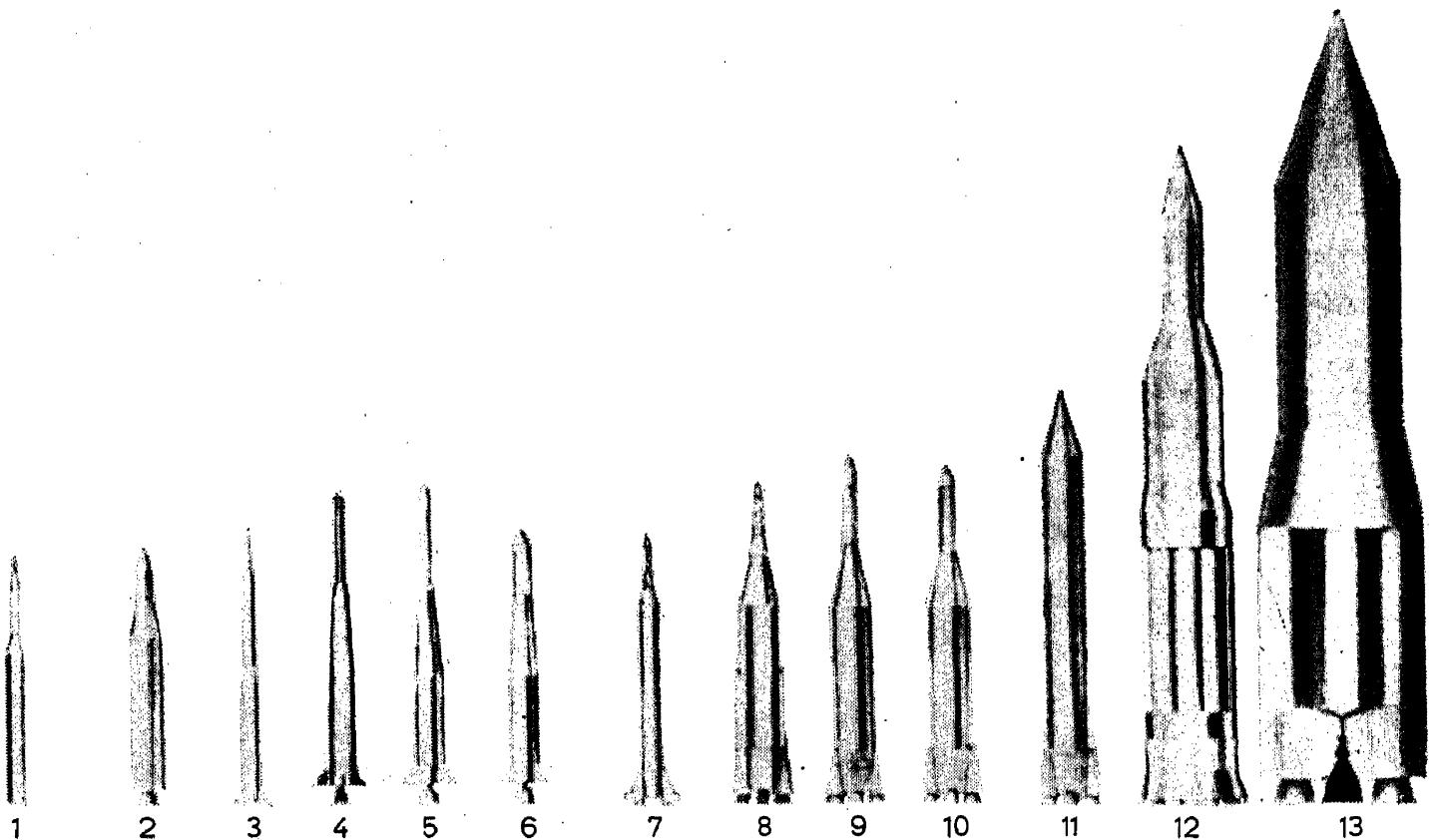


Fig. 7. - The U. S. spacecraft launch vehicles. 1) Vanguard; 2) Juno II; 3) Scout; 4) Thor-Able; 5) Delta; 6) Thor-Agena B; 7) Redstone; 8) Atlas; 9) Atlas-Able; 10) Atlas-Agena B; 11) Centaur; 12) Saturn C-1; 13) Nova.

return from its investment in terms of the obtaining of new scientific knowledge. This implies, for example, that back-up vehicles may not always be available in case of launch failures, since failure of a single spacecraft program is not catastrophic to the extent that the remaining space program will be held up appreciably by that failure, or that our national security is in jeopardy. On the other hand,

2) it is essential that NASA establish a high degree of reliability in the vehicle and spacecraft systems, since failure of one mission represents a loss of a significant fraction of the yearly effort. The NASA philosophy for obtaining high reliability can be stated as follows: *Reduce to a minimum the number of different types of vehicles, spacecraft, and components that are developed, and thereby increase the frequency with which those that remain are used.* NASA expects eventually to have available a fleet of standard vehicles and spacecraft whose reliability will more than offset the disadvantage of using vehicles and spacecraft on specific missions for which they may not be optimum.

Figure 7 shows the vehicles which have been used in the past or will be used in the future. Table I lists all spacecraft, both NASA and Department of Defense, which have been successfully launched by them, and indicates which vehicles will be used in the future for major programs now in process.

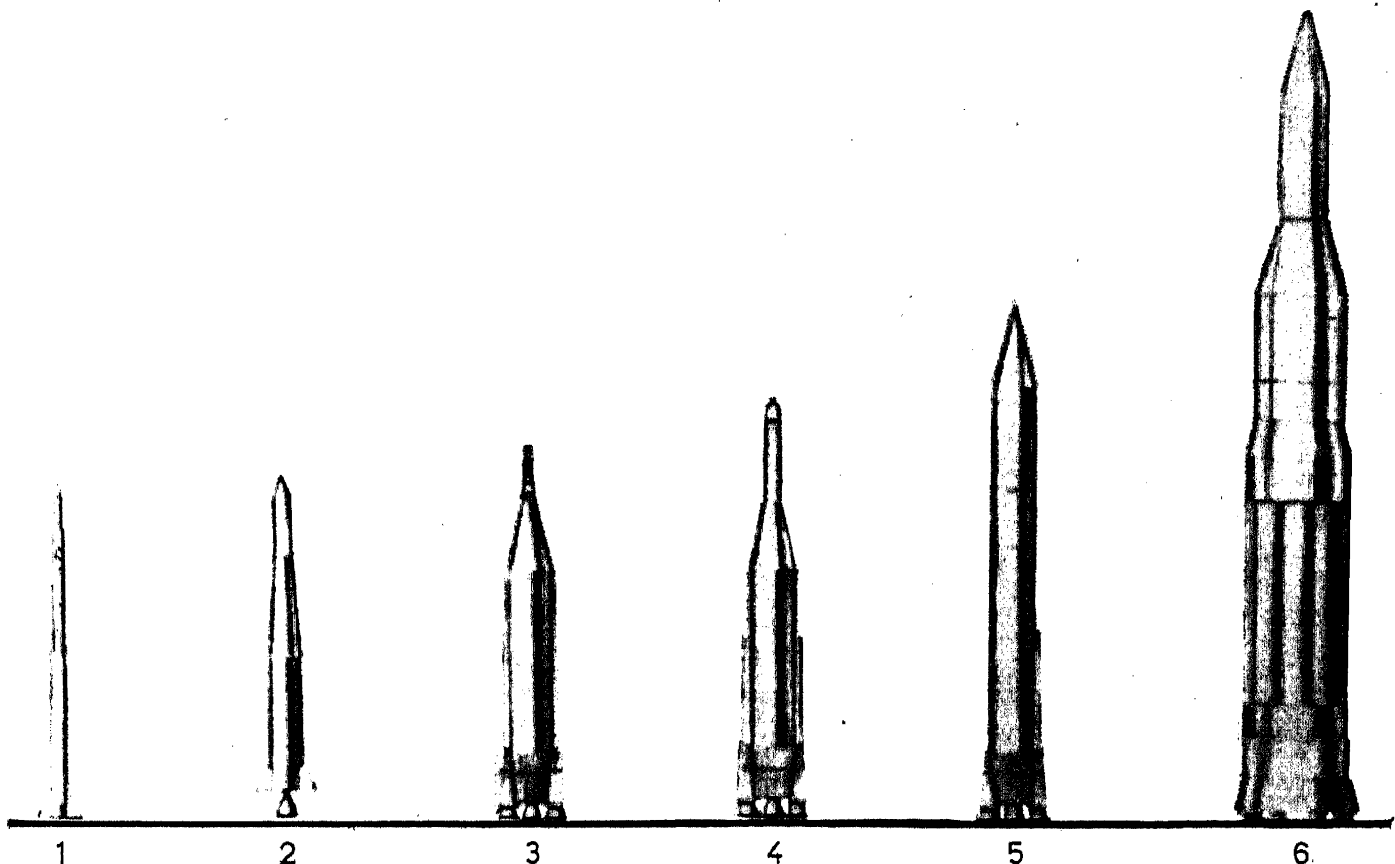


Fig. 8. - The NASA family of launch vehicles. 1) Scout; 2) Thor-Agena B; 3) Atlas; 4) Atlas-Agena B; 5) Centaur; 6) Saturn.

TABLE I. — *Launch vehicle summary.*

Vehicle	Stage	Propellant	Stage mass (kg)	Thrust (kg)	Max. Dia. (m)	Height Less Payload (m)	Payload (kg)		First successful launch	Spacecraft
							550 km Orbit	Escape		
Jupiter C	1	LOX/HYDYNE								
	2	SOLID	29 500	37 600	1.77	20.9	13.6	—	31 Jan 58	Explorer I, III, IV
	3	SOLID								
	4	SOLID								
Vanguard	1	LOX/KER	8 000	12 700 (SL)						
	2	UDMH/WIFNA	1 900	3 500	1.14	22.0	9.8 (Elliptic orbit)	—	17 Mar 58	Vanguard I, II, III
	3	SOLID	195	1 000						
Juno II	1	LOX/RP	54 400	68 000 (SL)						
	2	SOLID	340	6 800	2.68	23.5	43.1 (Max. load)	6.8	6 Dec 58	Pioneer III, IV, Explorer VII, VIII, XI
	3	SOLID	91	1 810						
	4	SOLID	41	725						
Thor-Able	1	LOX/RP	48 500	68 000 (SL)						
	2	WIFNA/UDMH	2 100	3 500	2.4	28.1	226	27.2	11 Oct. 58	Pioneer I, V, Explorer VI, Tiros I.
	3	SOLID	238	1 270						
Atlas-Able	1	LOX/RP		166 000 (SL)						
	2	LOX/RP	118 000	36 300	3.04	31.6	—	163	No successful launches	
	3	WIFNA/UDMH	2 100	3 500						
	4	SOLID	238	1 270						
Delta	1	LOX/RP	48 500	77 000						
	2	WIFNA/UDMH	2 100	3 500	2.4	27.0	363	59.0	13 May 60	Echo I, Tiros II, Explorer X
	3	SOLID	238	1 270						

	1	SOLID	10 700	46 700 (SL)								
Source	3	SOLID	1 220	6 170								
	4	SOLID	238	1 270								
Atlas	1	LOX/RP	118 000	166 000 (SL)	3.04	29.6	—	—	18 Dec 58	Score, Mercury (*)		
	2	LOX/RP										
Redstone	1	LOX/RP	30 000	35 400 (SL)	1.80	25.3	Ballistic	—	19 Dec 60	Mercury		
Atlas-Agena B	1	LOX/RP	118 000	166 000 (SL)	3.04	29.9	2 260	340	12 July 61	Ranger (*)		
	2	LOX/RP	6 350	36 300						OAQ (*), OGO (*)		
	3	IRFNA/UDMH		6 800						Samos II (**)		
Centaur	1	LOX/RP	118 000	166 000 (SL)	3.04	32.0	3 860	1 130	1963	Midas II (**)		
	2	LOX/RP	13 600	36 300						Mariner (*)		
	3	LOX/H ₂		13 600						OGO (*), Surveyer (*)		
Thor-Agena B	1	LOX/RP	48 500	75 000 (SL)	2.4	26.2	726	—	12 Nov 60	OSO (*), Nimbus (*), OGO (*), Discoverer (I, II, V, VI, VII, VIII, XI, XIII, XIV, XV, XVII, XVIII, XIX, XX, XXI, XXIII, XXV) (***)		
	2	IRFNA/UDMH	6 350	6 800								
Saturn (C-1)	1	LOX/RP	—	680 000 (SL)	6.58	45.7	8 620	2 260	1961			
	2	LOX/H ₂	—	31 800								
	3	LOX/H ₂	—	16 000								
Saturn (C-2)	1	LOX/RP	—	680 000 (SL)	6.58	—	20 400	—	1963			
	2	LOX/H ₂	—	363 000								
	3	LOX/H ₂	—	31 800								

(*) Future major program.

(**) Launched by earlier Atlas-Agena, now obsolete.

(***) Many of the earlier Discoverer spacecraft were launched by the earlier Thor-Agena, now obsolete.

If the philosophy of minimum variety is to be followed, then the number of different types of vehicles should be reduced. This, indeed, is the trend in the long-range program. Many of the vehicles shown are interim vehicles, chosen simply because they were most easily available. They have given good service, but will be phased out in favour of a set of five or six vehicles which will provide an adequate variety of payload capability.

Figure 8 shows the vehicles which NASA plans to retain in its vehicle program for the near future. The characteristics of these vehicles are as follows (*) (**).

Scout. - The Scout is a four stage, solid propellant vehicle (Fig. 9). The vehicle development program was initiated in late 1958 and the first successful

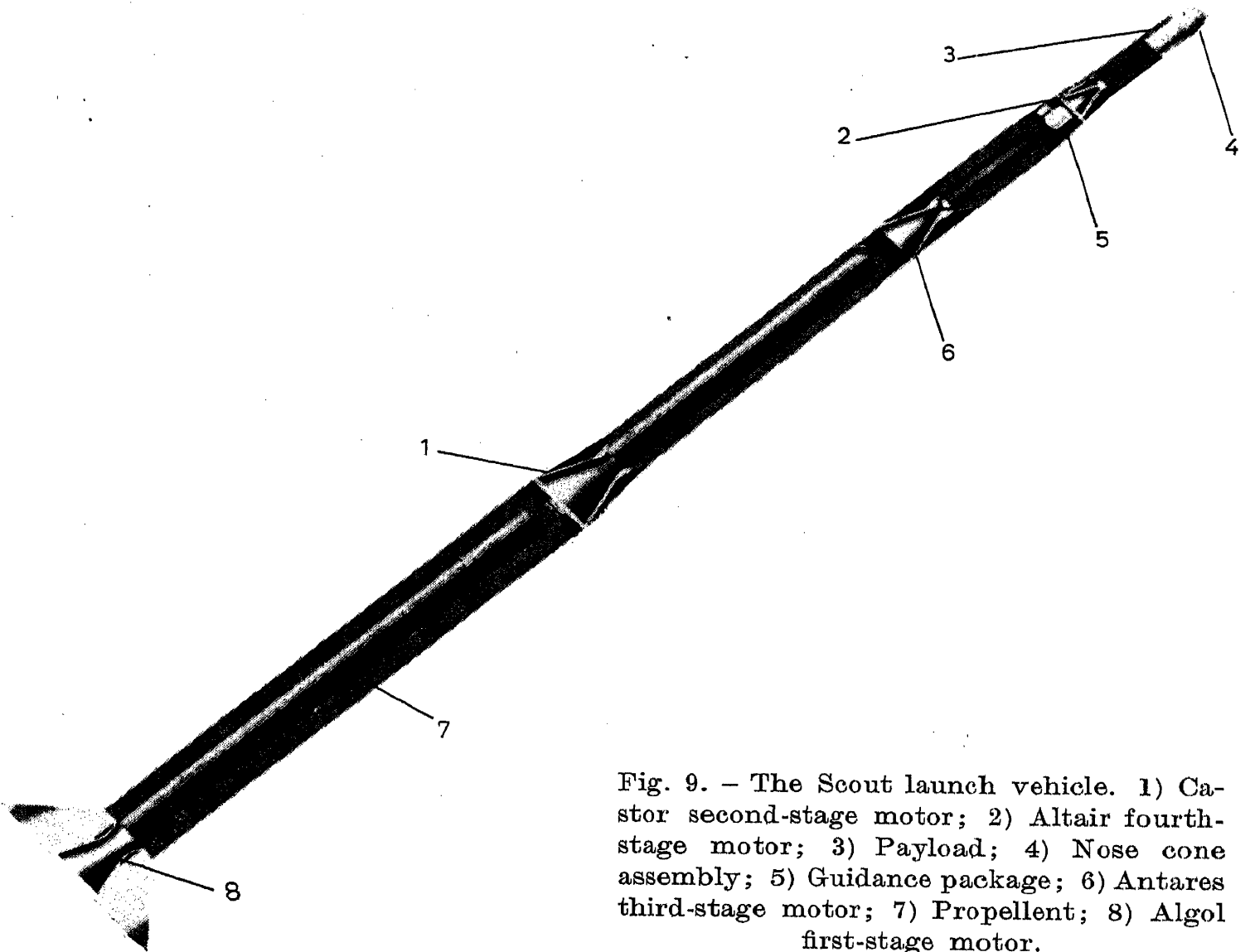


Fig. 9. - The Scout launch vehicle. 1) Castor second-stage motor; 2) Altair fourth-stage motor; 3) Payload; 4) Nose cone assembly; 5) Guidance package; 6) Antares third-stage motor; 7) Propellant; 8) Algol first-stage motor.

(*) Obtained from NASA-Industry Program Plans Conference, July 28-29 (1960).

(**) Note added in proof. - The delta vehicle has been added to the family of permanent vehicles.

launching of the complete vehicle occurred on 16 February 1961, when a 3.7 m diameter balloon satellite (Explorer IX) was orbited. The vehicle is expected to have a multiplicity of uses. As much as 10 kg can be placed into a 550 km circular orbit above the earth; vertical probes carrying useful payloads to heights as great as 7500 km can be launched. In addition, the Scout can be used for testing, within the atmosphere, very high speed air frame designs and re-entry bodies. The capability will exist for launching the Scout vehicle from both the Atlantic and the Pacific Missile Ranges as well as from Wallops Island, Virginia. One of the principal merits of the Scout is its relatively low cost. The cost of a production vehicle plus the cost of the launching operation may be less than one million dollars.

Thor-Agena B. - The Thor-Agena B, as the name implies, consists of a Thor intermediate range ballistic missile (IRBM) first stage and an Agena B second

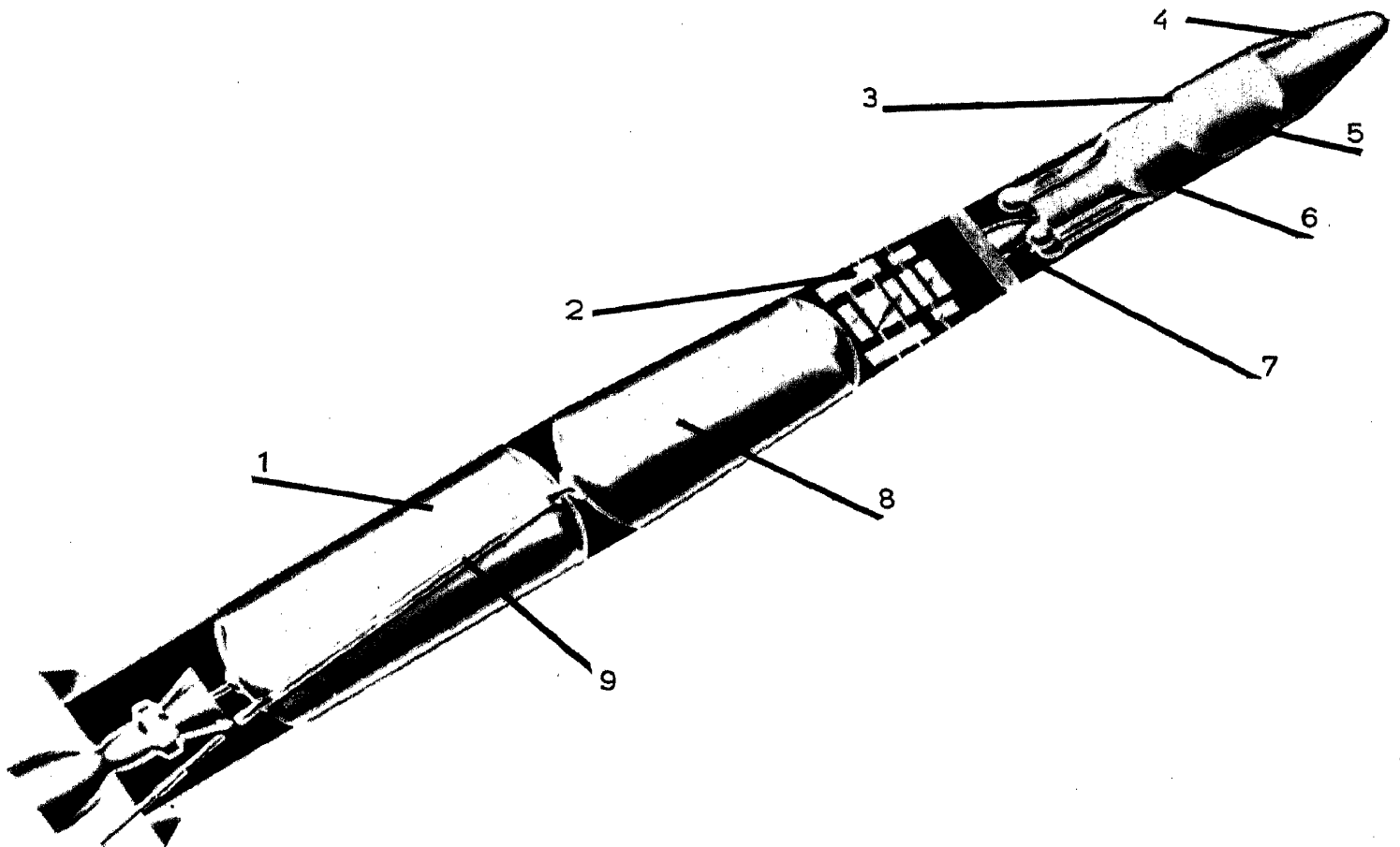


Fig. 10. - The Thor-Agena B launch vehicle. 1) Thor first-stage; 2) Guidance section; 3) Agena B second-stage; 4) Payload; 5) UDMH fuel; 6) IRFNA oxydizer; 7) Agena B engine; 8) RP-1 fuel; 9) LOX oxydizer.

stage (Fig. 10). This second stage is an enlarged version of the one that has been used successfully in the Air Force Discoverer Program. Because of the

relatively high vehicle reliability achieved in the Discoverer program, NASA elected to employ the Thor-Agena B as successor to the previous IRBM-based space vehicles such as the Thor-Able, and Juno II. It is planned that the Thor-Agena B will be used for launching both scientific and meteorological spacecraft. Present plans are that all of these vehicles will be launched at the Pacific Missile Range into highly inclined earth orbits.

Atlas-Agena B. — The Atlas-Agena B (Fig. 11) uses the same upper stage planned for the Thor-Agena B on an Atlas first stage. This vehicle will be used in a number of Air Force programs and the basic design will have under-

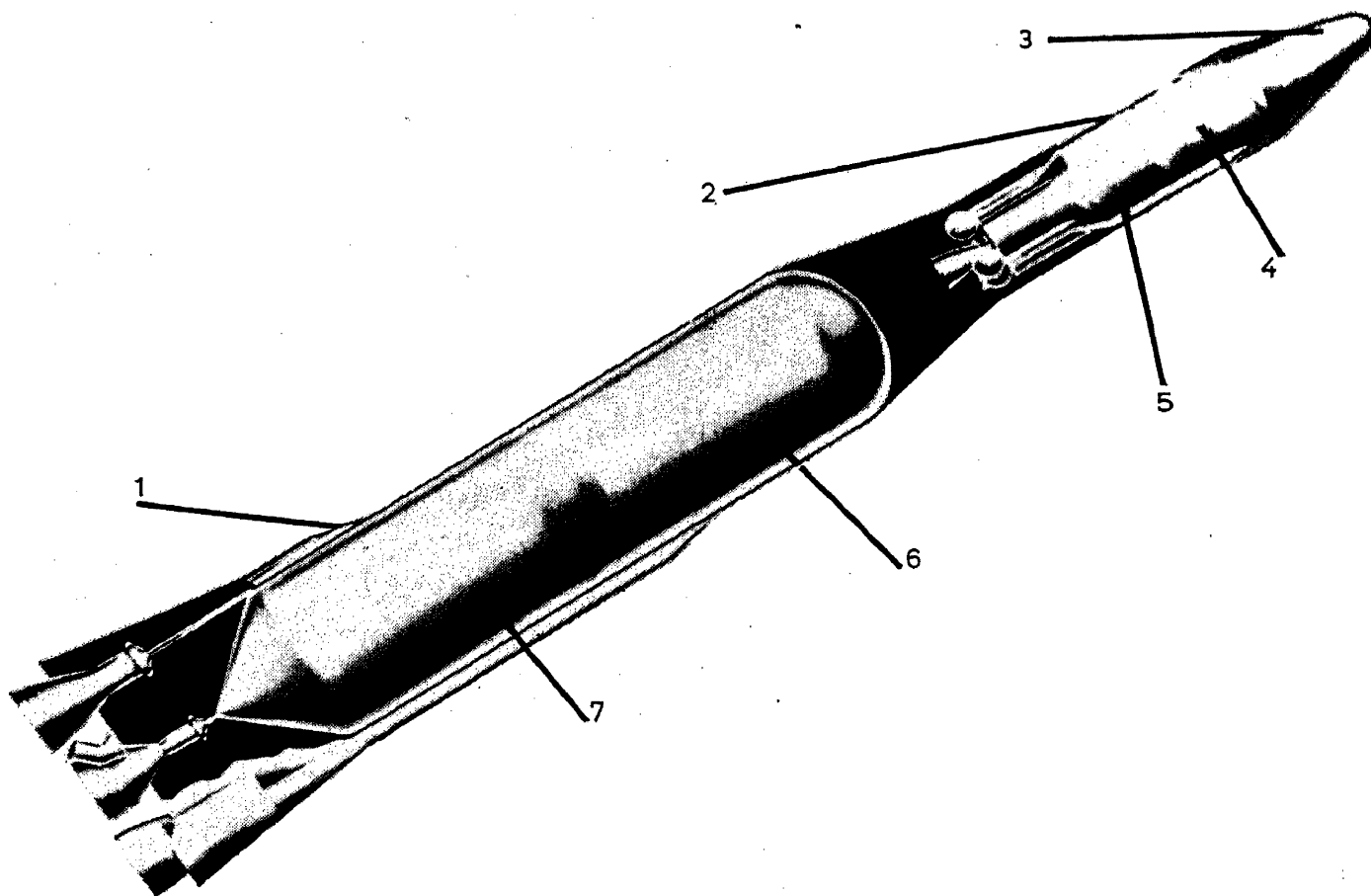


Fig. 11. — The Atlas-Agena B launch vehicle. 1) Atlas; 2) Agena B second-stage; 3) Payload; 4) UDMH fuel; 5) IRFNA oxidizer; 6) LOX oxidizer; 7) RP-1 fuel.

gone a number of firings before the first Atlas-Agena B is used for a NASA mission. NASA proposes to use the Atlas-Agena B principally for early lunar exploration and scientific earth satellites. In time, the Atlas-Agena B may be replaced by the Centaur. Present plans are to launch most of the Atlas-Agena B vehicles from the Atlantic Missile Range but a few will be launched from the Pacific Missile Range.

Centaur. — The Centaur vehicle consists of an Atlas first stage and a second stage which uses hydrogen and oxygen as propellents (Fig. 12). The high energy propellents give the Centaur the greatest capability of any Atlas-based

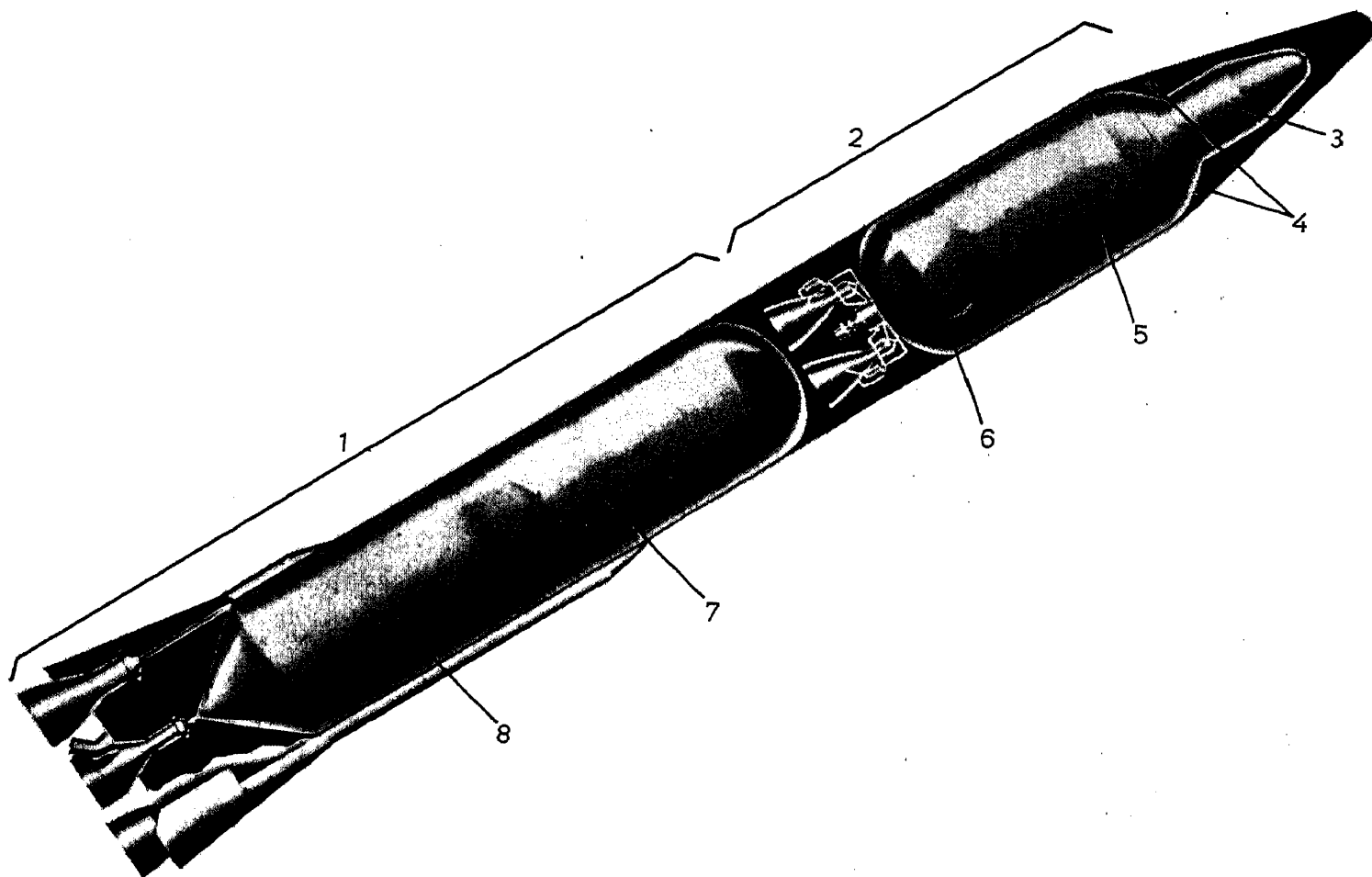


Fig. 12. — The Centaur launch vehicle. 1) Atlas first stage; 2) Second stage; 3) Payload; 4) Guidance package; 5) Liquid hydrogen tank; 6) LOX tank; 7) First stage LOX tank; 8) RP-1 fuel.

vehicle now programmed. Engineering design of the new engine is complete and one of these engines has been operated for more than 30 minutes at full thrust. The inertial guidance system for this vehicle has many advanced features. A great deal of research, engineering, and developmental testing yet remain to be done before the first launching attempt. Centaur will be employed for a variety of scientific missions including, for example, lunar and planetary exploration and geophysical studies. According to current plans, the Centaur vehicles will be launched from the Atlantic Missile Range.

Saturn. — Three versions of the Saturn vehicle group are being considered, namely, configurations C-1, C-2, and C-3. The C-1, with the least payload capability but earliest availability, is at present under active development (Fig. 13). It consists of a 680 000 kg thrust cluster of eight Rocketdyne

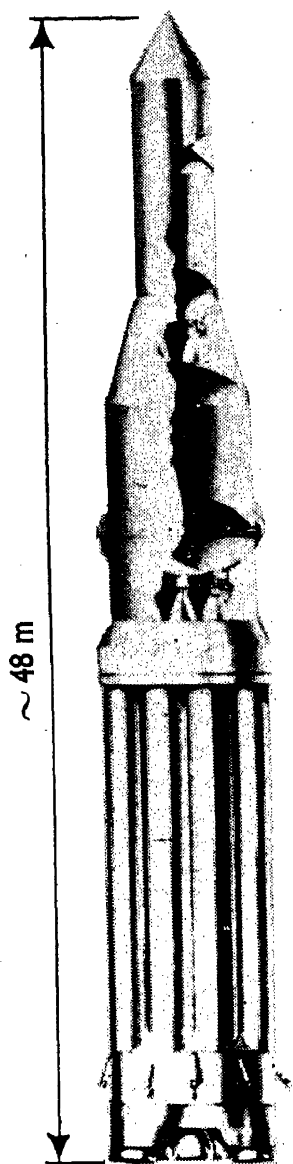


Fig. 13. - The Saturn C-1 launch vehicle.

engines as the booster, a second stage utilizing four Pratt and Whitney Centaur engines providing a total thrust of 31800 kg, and a third stage which is essentially the existing Centaur hydrogen-oxygen rocket modified as required for vehicle and spacecraft attachments. This Centaur engine will have a thrust of approximately 16 000 kg.

The C-2 is presently in the planning stages, parameter study, and early layout phase. It will utilize the same 680 000 kg thrust booster and upper stages as the C-1 vehicle but with an additional stage inserted between the booster and the C-1 second stage. This additional stage of the proposed C-2 vehicle will be a completely new development. The engines will be 91 000 kg thrust hydrogen-oxygen engines. Four of

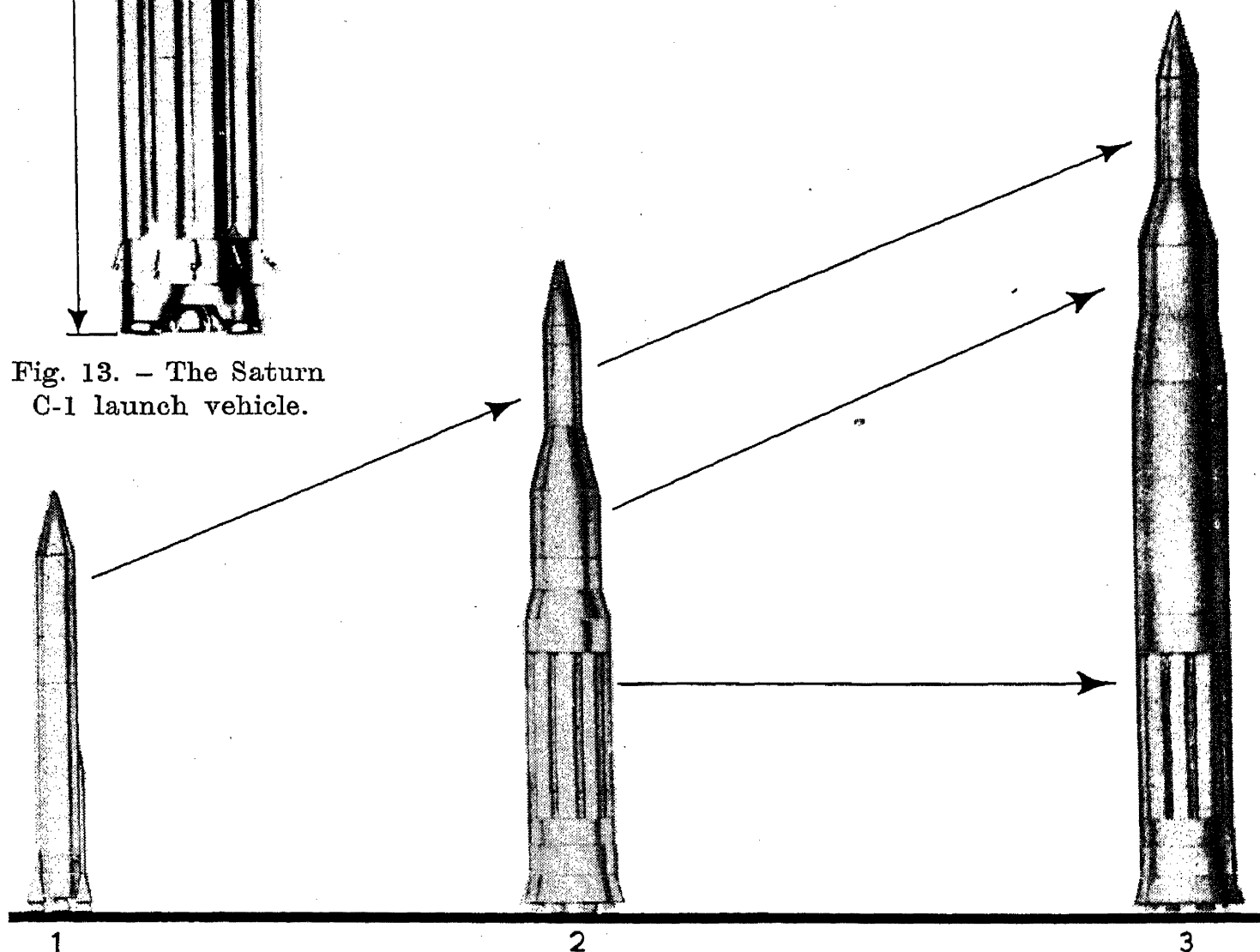


Fig. 14. - Drawing showing the Saturn C-2 and its evolution from the: 1) Centaur and 2) Saturn C-1 vehicles.

these engines will give the stage a thrust of approximately 360 000 kg. Figure 14 illustrates the manner in which the earlier vehicles will be incorporated into larger, more advanced vehicles, in keeping with the philosophy stated earlier.

The C-3 configuration, which would be the largest all-chemical version of the Saturn, has not gone beyond initial conceptual planning and performance calculations. Its development depends mainly on the results of extensive studies covering the full spectrum of space flight operations with the C-2 capability.

Present plans call for all Saturn C-1 and C-2 launchings to be made from the Atlantic Missile Range at Cape Canaveral.

It can be seen from the above summary that the NASA vehicle program includes a spectrum of launch vehicles from the small and relatively inexpensive Scout to the Saturn. Preliminary planning is underway for the development of the even larger Nova vehicle. It is expected that this selection of vehicles will be able to launch spacecraft having a wide range of configurations to accomplish a large variety of missions. The small vehicles will be used to launch relatively short lead-time payloads for conducting exploratory experiments and for the early investigation of new phenomena. The larger vehicles will allow the flight of a large number of related experiments for synoptic observations of the various phenomena.

5. - The spacecraft.

The philosophy outlined in the above discussion of the launch vehicle program applies also to the spacecraft program. The development of a relatively small number of observatory-type spacecraft which can be used over and over for performing many different experiments is expected to result in a greater reliability than the individual tailoring of the spacecraft for each new launching. This philosophy applies to the programs for the technological utilization of space and for manned space flight, as well as to the unmanned scientific research program.

The unmanned scientific program is divided into two major parts; experiments with rockets and satellites to investigate the earth and its environment, and experiments with space probes to explore the moon, the planets and interplanetary space. To accomplish the first part of this program, basic spacecraft systems and experiments are being designed for the Orbiting Solar Observatory (OSO), the Orbiting Geophysical Observatory (OGO) and the Orbiting Astronomical Observatory (OAO). It is expected that these observatories will be launched at regular intervals carrying whatever experiments are ready at that time, into orbits which are based on the requirements of the experiments

which are installed. The differences between these observatories are primarily differences between the stabilization and orientation systems. The Orbiting Solar Observatory, as its name implies, is designed to study the sun itself and

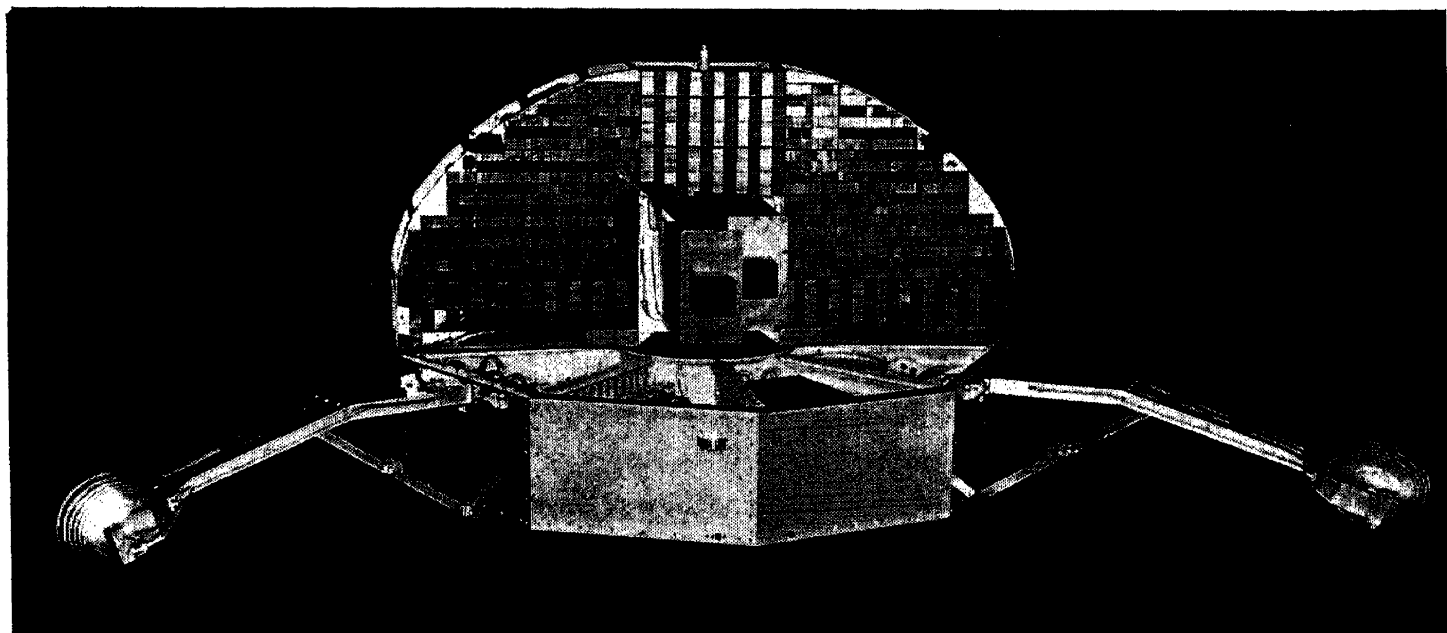


Fig. 15. — The Orbiting Solar Observatory (OSO).

phenomena directly related to the sun (see Fig. 15). Thus, its primary orientation will be toward the sun. This spacecraft is now under construction and will weigh about 160 kg when completely equipped with instruments. Basically, it consists of a flywheel-like section with attached arms which is spun to produce a large moment of inertia for gyroscopic rigidity, and a solar oriented section which contains the sensors and solar battery. A compressed gas jet system will keep the spin axis directed normal to the spacecraft-sun line, and a motor will produce a torque to maintain the rotation of the oriented section relative to the flywheel section. It is expected that orientation relative to the sun will be maintained with an accuracy better than a few minutes of arc. The first OSO launch is expected this year. One of the primary instruments in this spacecraft will be an ultra-violet spectrometer.

Figure 16 is an artist's conception of the Orbiting Astronomical Observatory which is expected to weigh about 1600 kg. The length of the central body will be about 3 m. This spacecraft will have a highly precise inertial stabilization and orientation system which will permit aiming the directional detectors in any arbitrary direction as directed by command. It will be capable of carrying a 0.9 m diameter telescope in the main body. The main spacecraft system will include the power, telemetry, stabilization, and the thermal control sub-systems which will be used by the appropriate experiments. It is expected that many types of astronomical experiments will be flown on the

OSO, including infra-red, ultra-violet, radio, and γ -ray experiments. Some of them are expected to employ photon-counting techniques to obtain very high sensitivities.

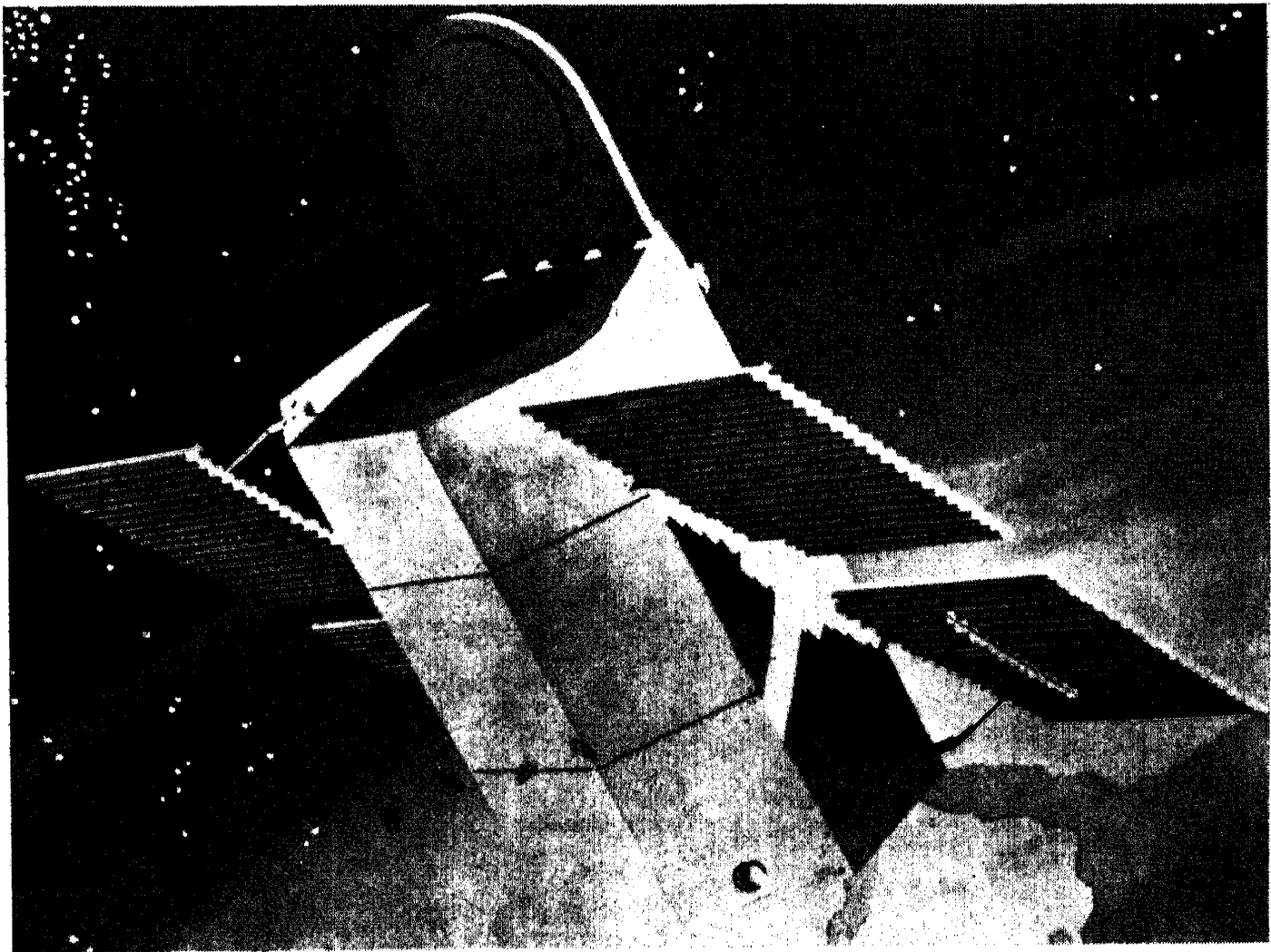


Fig. 16. — Artist's Conception of the Orbiting Astronomical Observatory (OAO).

The third and most universal type of large stabilized spacecraft system is the Orbiting Geophysical Observatory, which will carry a wide variety of experiments including those designed to investigate energetic particles, magnetic, electric, and gravitation fields, dust, atmospheric structure, the ionosphere, solar physics, astronomy, meteorology, and spacecraft technology. It is anticipated that this observatory will carry a larger proportion of particle and field experiments than will the two types previously described. Its primary orientation will be toward the earth but appendages will direct sensors toward the sun and in the plane of the orbit. Figure 17 is an artist's drawing of the OGO configuration. It will weigh a total of about 450 kg, of which at least 68 kg will be reserved for the experiments themselves. The standard observatory will include the power, data handling, temperature control, and stabi-

lization subsystems which will remain essentially unchanged from mission to mission. A number of boom-like extensions will be provided to support experiments which are sensitive to the spacecraft induced environmental disturbances. It will be possible to launch the OGO spacecraft into a wide variety of orbits ranging from a low altitude (270 km perigee near-circular) polar orbit with the use of the Thor-Agena B launch vehicle, to a high eccentricity (270 km perigee to 110 000 km apogee) 31° inclination orbit by the use of the Atlas-Agena B vehicle. It is expected that the first Eccentric Orbiting Geophysical Observatory (EOGO) will be launched in mid-1963, and that the first Polar Orbiting Geophysical Observatory (POGO) will be launched in mid-1964.

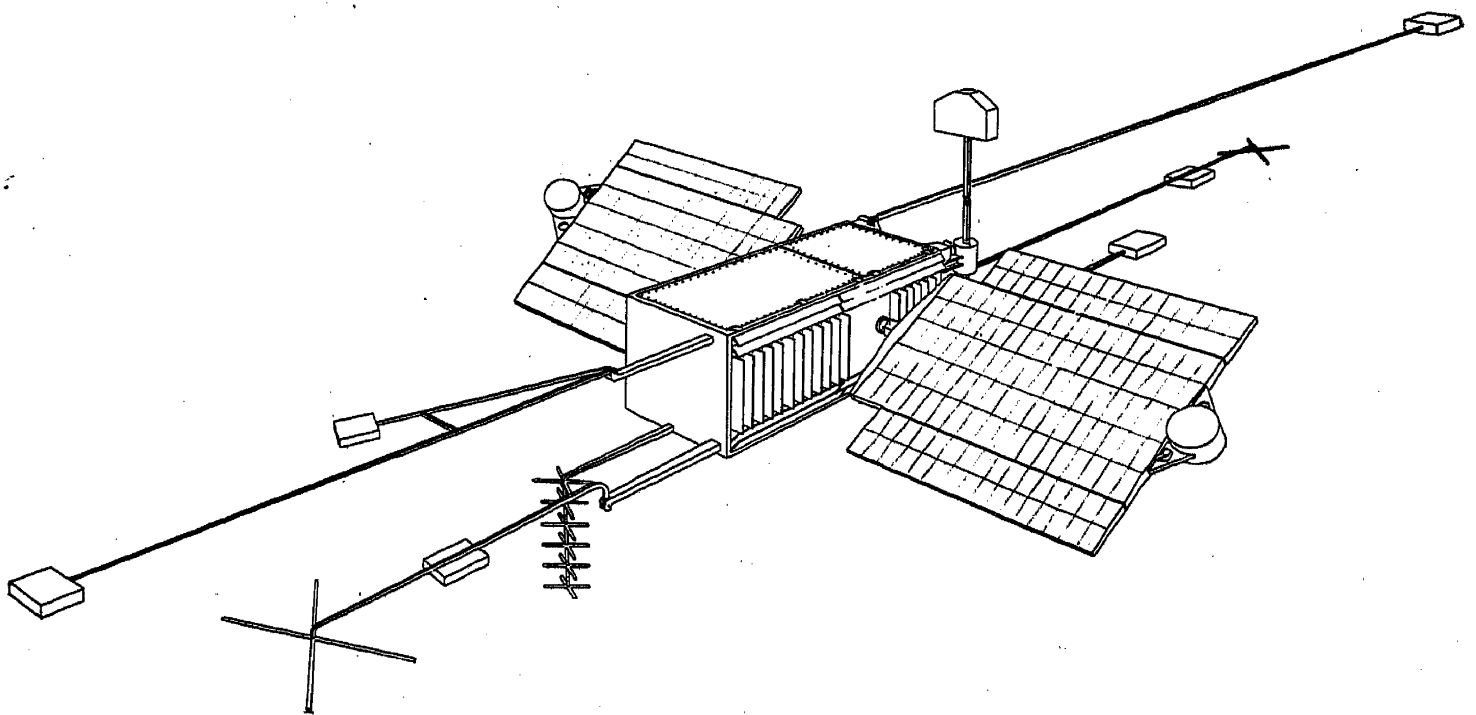


Fig. 17. - Drawing of the Orbiting Geophysical Observatory (OGO).

Although an effort is being made to avoid custom design and construction for each satellite by the use of the standardized observatories, it is expected that there will continue to be a need for some satellites which are tailored to small groups of specific scientific experiments. This will be true particularly for the small Scout spacecraft on which the weight penalty for standardized design is not acceptable. These small spacecraft will be used extensively in the very near future until the heavier and more complex observatories described above become available.

Experiments which need to spend a considerable period of time in interplanetary space, or which are intended to investigate the moon and planets, will be carried on one of a series of interplanetary spacecraft. The first of these is Ranger, depicted in Fig. 18, intended for use in lunar missions with

gross weights varying from 300 to 550 kg. The first two of this series will carry loads of predominately energetic particles and magnetic field experiments to a radial distance of 10^6 km from the earth. The later lunar impact missions will use a modified version of the basic spacecraft, which will carry a serviceable capsule containing a seismometer and experiments to investigate the moon's surface. During the early stages of these flights from the earth to the moon,

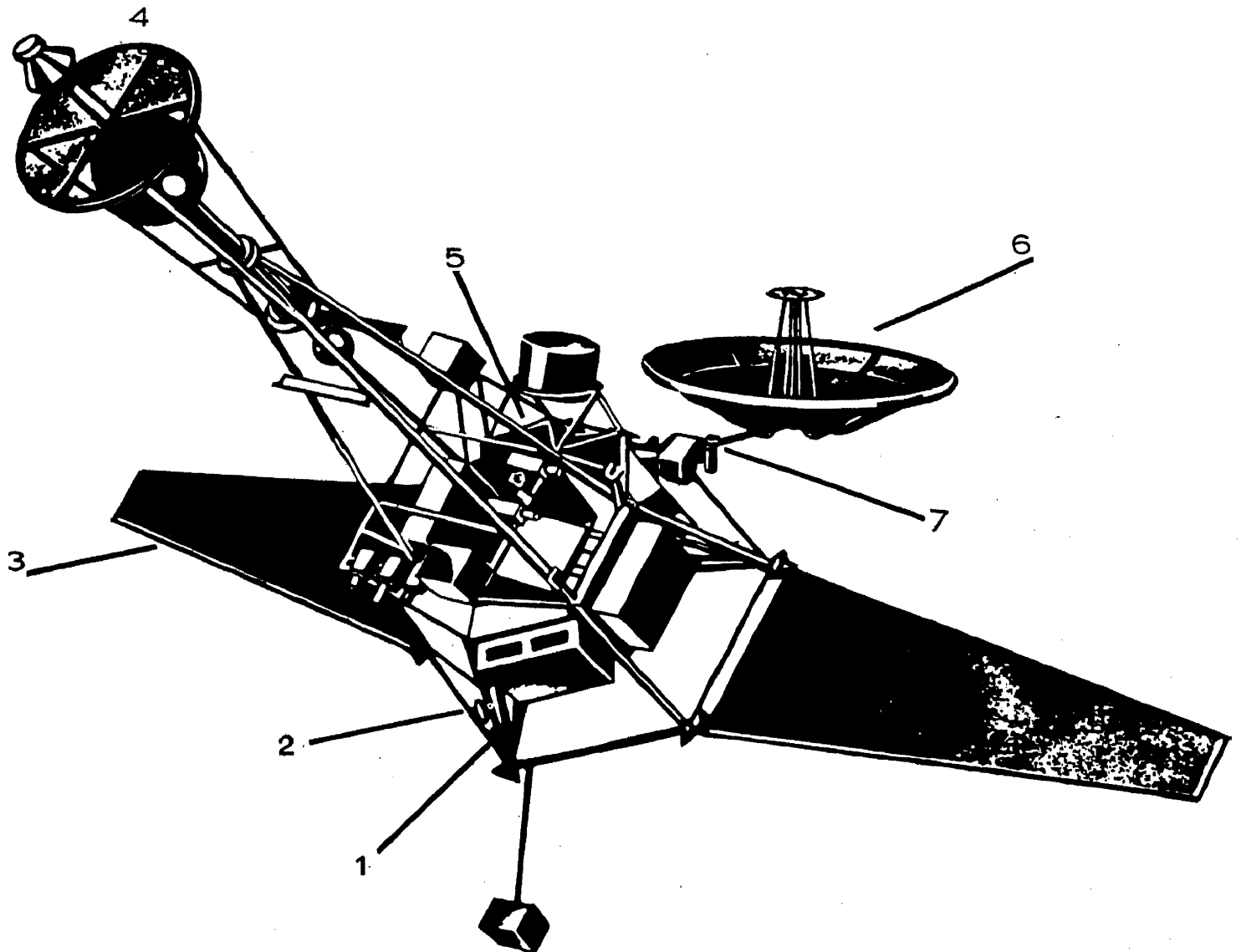


Fig. 18. -- The Ranger lunar spacecraft.

the Ranger spacecraft will maintain three-axis attitude control with its antenna pointing toward the earth and its solar panels directed toward the sun. Radio tracking will reveal any necessary course corrections, which the spacecraft will make on command by orienting itself for a mid-course rocket firing, and then reorienting itself as before. As the spacecraft approaches the lunar surface, it will reorient itself on command to direct its major axis parallel to the vertical descent path, and will then begin taking high resolution television pictures and performing experiments intended to find out more about the lunar

surface. At a low altitude, the survivable capsule will be slowed by a retro-rocket for a rough but safe landing on the moon's surface, where it will perform other experiments.

Very closely related to the Ranger configuration is the Mariner spacecraft, which is being developed for interplanetary flights to Venus and Mars. The basic spacecraft frame is nearly identical with that of the Ranger spacecraft. The solar panels are larger and additional equipment has been included to make possible transmission from the larger distances involved. The first of these missions is to Venus and will include a large variety of energetic particles and magnetic field experiments intended to investigate the entire charged particle energy spectrum from near thermal to high energy cosmic rays. A later mission will carry experiments to the near vicinity of the planet Mars.

The next program in this series will utilize the Surveyor spacecraft which will be intended to soft-land experiments on the lunar surface. It will be used primarily for analysis of the structure of the moon and the properties of its surface.

6. - Conclusion.

This paper has outlined the United States NASA particles and fields space-research program and has discussed the launching vehicles and the spacecraft intended to be used for many different orbits and groupings of experiments. It is believed that this program will provide, within the next few years, the solutions to some of the presently existing questions and the cration of many more questions which will need to be answered in our search for a better understanding of our physical universe.

REFERENCES

- [1] R. H. GODDARD: *A Method of Reaching Extreme Altitudes*, Smithsonian, Misc. Collections, vol. 71 (1919).
- [2] J. A. VAN ALLEN, G. H. LUDWIG, E. C. RAY and C. E. MCILWAIN: *Jet Propulsion*, 28, 588 (1958).
- [3] J. A. VAN ALLEN, C. E. MCILWAIN and G. H. LUDWIG: *Journ. Geophys. Res.*, 64, 271 (1959).
- [4] J. A. VAN ALLEN, C. E. MCILWAIN and G. H. LUDWIG: *Journ. Geophys. Res.*, 64, 877 (1959).
- [5] P. ROTHWELL and C. E. MCILWAIN: *Nature (London)*, 184, 138 (1959).
- [6] P. ROTHWELL and C. E. MCILWAIN: *Magnetic Storms and the Van Allen Radiation Belts: Observations from Satellite 1958 Epsilon (Explorer IV)*, in *Space Research*, edited by H. KALLMAN BIJL (Amsterdam, 1960), pp. 897-909.

- [7] P. ROTHWELL and C. E. MCILWAIN: *Journ. Geophys. Res.*, **65**, 799-806 (1960).
- [8] J. A. VAN ALLEN and L. A. FRANK: *Nature (London)*, **183**, 430 (1959).
- [9] J. A. VAN ALLEN and L. A. FRANK: *Nature (London)*, **184**, 219 (1959).
- [10] J. A. VAN ALLEN: *Journ. Geophys. Res.*, **64**, 1683 (1959).
- [11] A. ROSEN, C. P. SONETT, P. J. COLEMAN, Jr. and C. E. MCILWAIN: *Journ. Geophys. Res.*, **64**, 709 (1959).
- [12] R. L. ARNOLDY, R. A. HOFFMAN and J. R. WINKLER: *Measurements of Van Allen Radiation Belts during Geomagnetic Storms*, in *Space Research*, edited by H. KALLMAN BIJL (Amsterdam, 1960), pp. 877-896.
- [13] R. L. ARNOLDY, R. A. HOFFMAN and J. R. WINKLER: *Journ. Geophys. Res.*, **65**, 1361 (1960).
- [14] C. Y. FAN, P. MEYER and J. A. SIMPSON: *Trapped and Cosmic Radiation Measurements from Explorer VI*, in *Space Research*, edited by H. KALLMAN BIJL (Amsterdam, 1960), pp. 951-966.
- [15] A. ROSEN, T. A. FARLEY and C. P. SONETT: *Soft Radiation Measurements on Explorer VI Earth Satellite*, in *Space Research*, edited by H. KALLMAN BIJL (Amsterdam, 1960), pp. 938-950.
- [16] C. P. SONETT, E. J. SMITH, D. L. JUDGE and P. J. COLEMAN, Jr.: *Phys. Rev. Lett.*, **4**, 161 (1960).
- [17] C. P. SONETT, E. J. SMITH and A. R. SIMS: *Surveys of the Distant Geomagnetic Field: Pioneer I and Explorer VI*, in *Space Research*, edited by H. KALLMAN BIJL (Amsterdam, 1960), pp. 921-937.
- [18] C. E. MCILWAIN and P. ROTHWELL: *Bull. Am. Phys. Soc.*, **4**, 238 (1959).
- [19] C. E. MCILWAIN and P. ROTHWELL: *Journ. Geophys. Res.*, **65**, 2508 (1960).
- [20] B. J. O'BRIEN, J. A. VAN ALLEN, F. E. ROACH and C. W. GARTLEIN: *Journ. Geophys. Res.*, **65**, 2759 (1960).
- [21] B. J. O'BRIEN and G. H. LUDWIG: *Journ. Geophys. Res.*, **65**, 2695 (1960).
- [22] J. A. VAN ALLEN and W. C. LIN: *Journ. Geophys. Res.*, **65**, 2998 (1960).
- [23] P. J. COLEMAN, Jr., L. DAVIS and C. P. SONETT: *Phys. Rev. Lett.*, **5**, 43 (1960).
- [24] C. Y. FAN, P. MEYER and J. A. SIMPSON: *Journ. Geophys. Res.*, **65**, 1862 (1960).
- [24] C. Y. FAN, P. MEYER and J. A. SIMPSON: *Phys. Rev. Lett.*, **5**, 269 (1960).
- [26] C. Y. FAN, P. MEYER and J. A. SIMPSON: *Phys. Rev. Lett.*, **5**, 272 (1960).
- [27] R. L. ARNOLDY, R. A. HOFFMAN and J. R. WINKLER: *Journ. Geophys. Res.*, **65**, 3004 (1960).
- [28] P. J. COLEMAN, Jr., C. P. SONETT, D. L. JUDGE and E. J. SMITH: *Journ. Geophys. Res.*, **65**, 1856 (1960).
- [29] C. P. SONETT, D. L. JUDGE, A. R. SIMS and J. M. KELSO: *Journ. Geophys. Res.*, **65**, 55 (1960).
- [30] J. E. NAUGLE and D. A. KNIFFEN: *Phys. Rev. Lett.*, **7**, 3-6 (1961).
- [31] L. R. DAVIS, C. E. FICHEL, D. E. GUSS and K. W. OGILVIE: *Phys. Rev. Lett.*, **6**, 492-494 (1961).

Spacecraft Information Systems.

G. H. LUDWIG

*Goddard Space-Flight Center, National Aeronautics and Space Administration
Greenbelt, Md.*

1. - Introduction.

The value of earth satellites and space probes for the gathering of scientific information about our physical universe is well established. Since these spacecraft are very expensive, it is necessary that we give a considerable amount of thought to the problem of maximizing the amount of information which can be obtained from each one. This paper contains a discussion of some of the difficulties involved in conducting an experiment in space, and of some of the techniques which are available for this purpose.

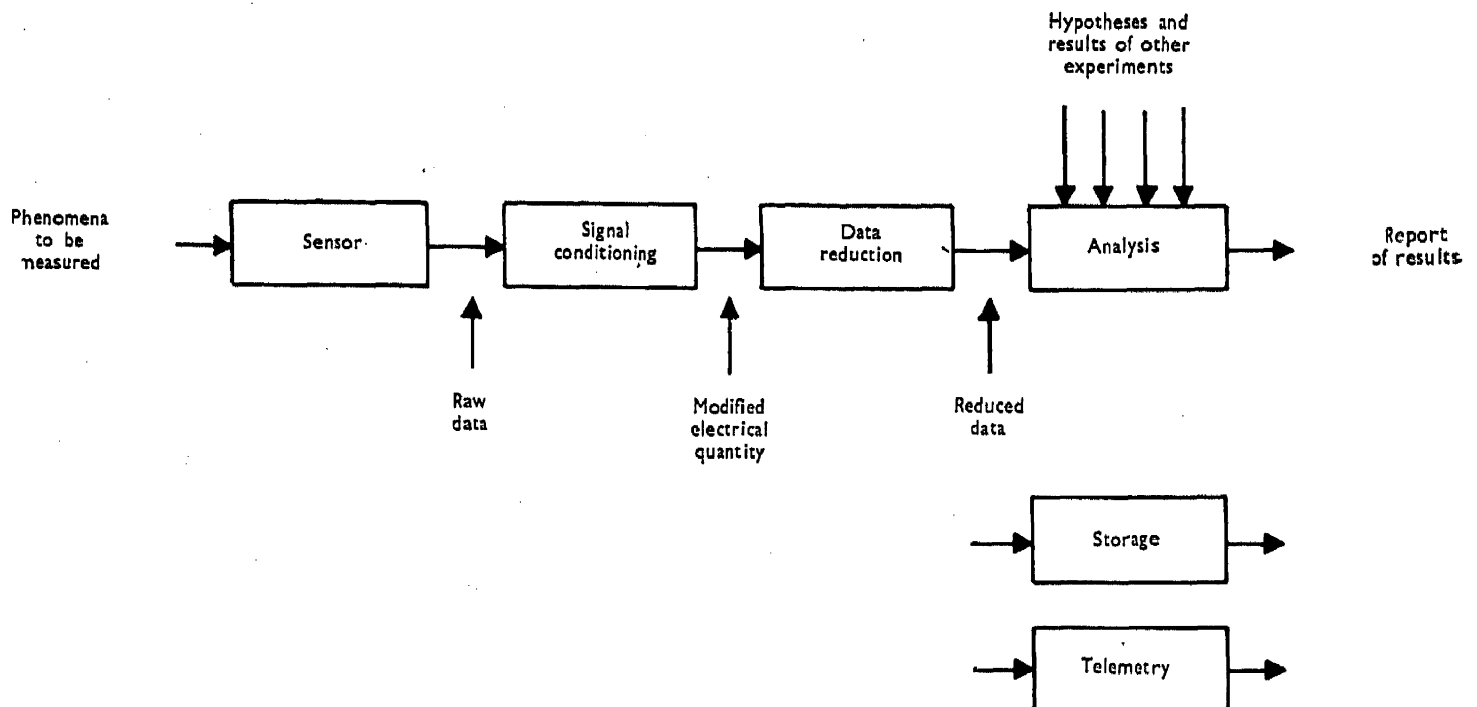


Fig. 1. - Experiment information flow.

We may think of any experiment, no matter where it is conducted, as a generalized system, as shown in Fig. 1. The end elements are the basic sen-

sors and the final report of results. The intermediate elements are the signal conditioning instruments which convert the signals from these sensors into usable electrical quantities, the reduction of these electrical quantities into some form which can be used by the experimenter, and the analysis of these data. In addition, for many experiments it is necessary to add blocks representing storage in one or in several places. Storage may actually be a part of all the blocks listed in Fig. 1. For example, if the sensor is a self-integrating ion chamber of the Neher type, it provides a storage function of its own. If the pulses which appear at the output of this chamber are counted, then this scaler represents a second form of storage. A tape recorder is often incorporated into the signal conditioning and data reduction systems; the placing of the reduced data on an oscillograph record or in tabular form is also storage.

The major difference between the general outline shown above and a spacecraft system is the addition of a block representing telemetry of the data from the spacecraft to the earth. This block can be placed anywhere along the line of data flow. In general, the information bit rate along this line of data flow decreases progressively as one moves from left to right in the diagram. It must be remembered that the experimenter first sees his data at the output of the telemetry system. The factors which determine where along the line of data flow the telemetry system should be placed for any specific experiment include:

- 1) The ability of the experimenter to predict the results. If he knows where in the possible range of values his results will occur he may perform a considerable amount of processing in the satellite before he sees the data. If, for example, he is measuring the energy spectrum of protons in the cosmic ray flux, and if he already has a *general* idea of the energy spectra of all the constituents, he might concentrate on the protons and reject all other particles. He might go so far as to include a computer in the spacecraft which would determine an analytic function which would represent the spectrum. In this case he would need to telemeter only a set of coefficients rather than an extensive set of data points. On the other hand, if he knows very little about the phenomenon, it may be necessary for him to telemeter all data directly from the sensor, and then to determine how the data should be conditioned and reduced on the ground after he has completed a preliminary investigation of the raw data.

- 2) Information bandwidth available for data transmission. Obviously, the first alternative listed above would require far less information bandwidth than the second.

- 3) System reliability. The larger the amount of data processing in the spacecraft, the greater is the complexity and the smaller is the system reli-

TABLE I. — *Energetic charged particle detectors.*

Detector type	Physical process utilized	Form of output	Characteristics of output utilized	Use in U.S. satellites
Ion chamber	Ionization of gas	a) Pulse b) Integrated current	a) Pulse rate b) Charge per pulse c) Charge per unit time	Used in Pioneer I, V, Explorer VI, VII
Proportional counter	Ionization of a gas and ion multiplication	Pulse	a) Pulse rate b) Charge per pulse	Used in Pioneer V, Explorer VI
Geiger Müller counter	Ionization of a gas and ion multiplication	a) Pulse b) Integrated current	a) Pulse rate b) Charge per unit time	Used in Explorer I, III, IV, VI, VII, Pioneer I, III, IV, V
Cloud chamber	Ionization and condensation in a supersaturated gas	Visual tracks recorded by a photographic film or television system	a) Rate of events b) Ionization density along track c) Curvature of track when magnetic or electric fields are applied d) Study of sequence of interactions and secondary particles produced	Not presently feasible due to need for recovery or transmission of high resolution pictures
Bubble chamber	Ionization and vaporization in a superheated liquid	Same as a cloud chamber	Same as a cloud chamber	Same as a cloud chamber
Solid	Solid scintillator followed by image intensifier and picture system	Same as a cloud chamber	Same as a cloud chamber	Same as a cloud chamber

NUCLEAR emulsions	IONIZATION FROM PARTICLES PRODUCES DEVELOPABLE SILVER HALIDE GRAINS	VISUAL RECORDS IN FORM OF EMULSION USUALLY OBSERVED UNDER MICROSCOPE	NUMBER OF MEASUREMENTS PER	VEHICLES CURRENTLY IN USE
Ion collector	Direct interception of low energy charged particles	a) Pulse b) Integrated current	a) Pulse rate b) Charge per pulse c) Charge per unit time	Used in Explorer VIII, X
Scintillation detector	Ionization or excitation of atoms in a gas, liquid or solid and production of photons during de-excitation. Detection of photons by light-sensitive detector such as photomultiplier tube	a) Pulse b) Integrated current	a) Pulse rate b) P. M. charge per pulse c) P. M. charge per unit time d) Pulse shape	Used in Explorer IV, VI, XI, Pioneer V
Čerenkov detector	Čerenkov radiation in a gas, liquid, or solid. Detection of the photons by light sensitive detector such as photomultiplier tube	a) Pulse	a) Pulse rate b) P. M. charge per pulse c) Direction of propagation of photons	Used in Explorer XI
Solid state detector	a) Ionization of atoms in a solid with resultant change of body conductivity b) Ionization of atoms in a solid and transport of resultant electrons across a potential barrier (p-n junction)	a) Resistance b) Pulse	a) Conductivity b) Pulse rate c) Voltage amplitude of pulse	To be used in near future

very used in SBE, NERV, Atlas pod rocket firings

bility. If the experimenter were to telemeter only the coefficients mentioned above, he would need to have a high degree of confidence in his instrumentation.

4) Time available for the preparation of the experiments. Complex spacecraft data processing systems demand a long time for their development and calibration.

Obviously, the choice of best location along the line of information flow for placement of the telemetry system can be expected to vary widely, even between experiments on the same spacecraft. Thus, any acceptable central spacecraft data handling system must be carefully designed to permit the necessary flexibility.

2. - General considerations.

On spacecraft containing more than one experiment the signal conditioning instrumentation is often split into two sections. The first consists of the detectors and the circuits which process their outputs up to the point where the central data handling system combines the data from a number of experiments. The second includes the instrumentation in this central system, which may further process and store the data before transmission. As the attempt is made to measure the characteristics of the phenomena in more and more detail, more and more processing of the data in the spacecraft before transmission will become necessary. This trend is expected to continue to the point that programmable computers will be used aboard the spacecraft to perform complex analyses. This tendency is a result of the fact that less payload weight will be required for the computers than would be required to produce the electrical power to telemeter the very large quantity of raw data, and that conservation of ground receiving station capability and data reduction facilities will be necessary.

A large number of different detectors exist for the study of energetic particles. Table I lists these detectors and indicates their applicability for use on spacecraft. Most of the detectors rely on the process of ionization for operation. The particle being detected produces ionization in the detector, and the electrons, ions, and photons which result are gathered and analysed. Some of the characteristics of the electrical signals which appear at the outputs of the detectors are summarized in Table I. These are the signals which must be prepared for storage and transmission over the telemetry link by the signal conditioning instrumentation.

The raw information must be converted into a form which meets a number of conditions:

- 1) It must be in the most meaningful form practicable.
- 2) It must be capable of being transmitted over a telemetry link.
- 3) The telemetry must be accomplished within the smallest information bandwidth practicable.
- 4) The instrumentation must meet the requirements of ruggedness, relatively low operating power, light weight, and relatively small volume imposed by the launching vehicle capabilities and the launching environment.
- 5) The instrumentation must be reliable while operating unattended for long periods of time in the space environment.

The operation most often performed on the outputs of those detectors which are basically electronic in nature include logic operations (including coincidence, anti-coincidence, etc.) pulse shaping, impedance transformation, counting of events, measuring of time intervals, sorting, amplification (either linear or non-linear and of either pulses or analog quantities), conversion from analog to digital form, and storage.

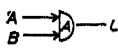
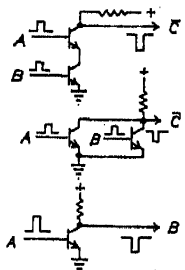
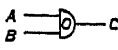
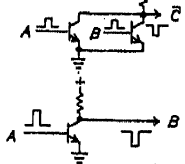
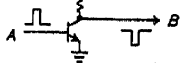
3. - Basic circuits.

Many experiments use a number of individual detectors in arrays. The outputs of these detectors are often acted upon by logic circuits. It is convenient to represent the logic operations by a special symbolic language [1]. In this symbolization, $AB = C$ represents the logic operation «output C if A and B », $A + B = C$ represents the operation «output C if A or B », and $\bar{A} = B$ represents the operation « B if not A ». The operations AB and $A + B$ are accomplished by use of «and» and «or» logic circuits, as shown in Table II. Commutative, associative, and distributive laws hold for these operations as follows:

Commutative laws:	$AB = BA$
	$A + B = B + A$
Associative laws:	$A(BC) = (AB)C$
	$A + (B + C) = (A + B) + C$
Distributive laws:	$A(B + C) = AB + AC$
	$A + (BC) = (A + B)(A + C)$

By the use of these laws and Table II, it is possible to describe any logic system in terms of a symbolic expression.

TABLE II. - *Logic Operations*

Symbolic	Description	Block diagram	Schematic diagram
$AB + C$	C if A and B		
$A + B = C$	C if A or B		
$\bar{A} = B$	B if not A		

A number of the detectors listed in Table I have pulse outputs which can be acted upon by these logic systems and by scalers with a minimum of preconditioning. For example, those detectors from which pulses having approximately fixed characteristics result from any detectable event, such as G.M. counters, utilize nonlinear amplifier and pulse-shaping circuits to provide proper driving signals for later circuits, and to assist in establishing the threshold characteristics for the system. Figure 2 shows a simple pulse amplifier and shaping circuit used with G.M. counters to produce driving pulses for scaling circuits. It is a simple current-fed driver which produces a saturating output pulse when the G.M. tube discharge current passes through the base-emitter junction. This circuit is very simple and contains a small number of com-

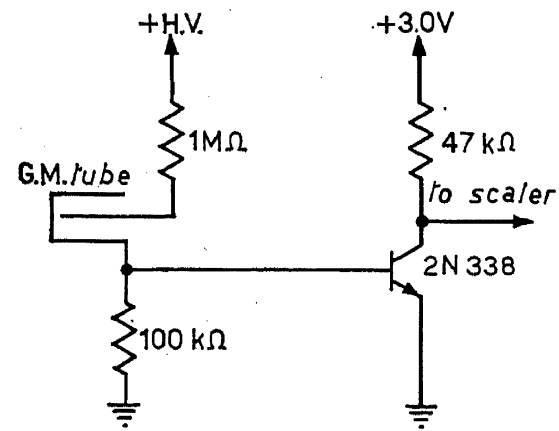


Fig. 2. - Simple G.M. tube driver (Explorer I, III).

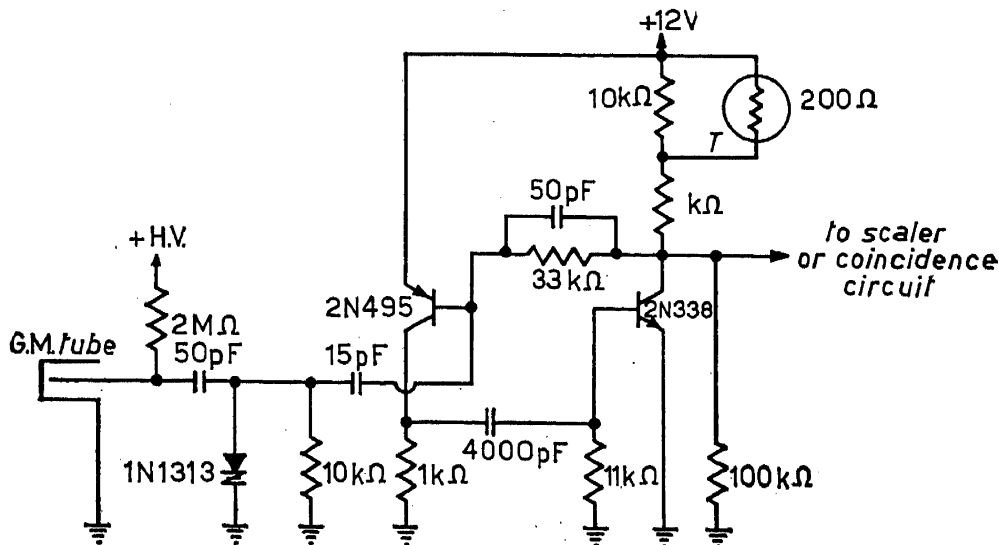


Fig. 3. - Monostable multivibrator G.M. tube driver (S-3).

ponents, but has the disadvantage that the system deadtime characteristics depend on counter pulse-amplitude, deadtime characteristics and the scaler input pulse amplitude threshold, which both depend to some extent on the supply voltages and operating temperature. The circuit of Fig. 3 is a partial solution to this problem. It utilizes the voltage pulse obtained from the G.M. tube center wire to trigger a monostable multivibrator. Thus, the total dead time, is, to a first approximation, the sum of the multivibrator pulse width and recovery time. It is interesting to note that in this multivibrator circuit essentially zero current flows except when a pulse occurs, due to the fact that both transistors in the multivibrator are cut off in the quiescent condition. They conduct only during the pulse interval.

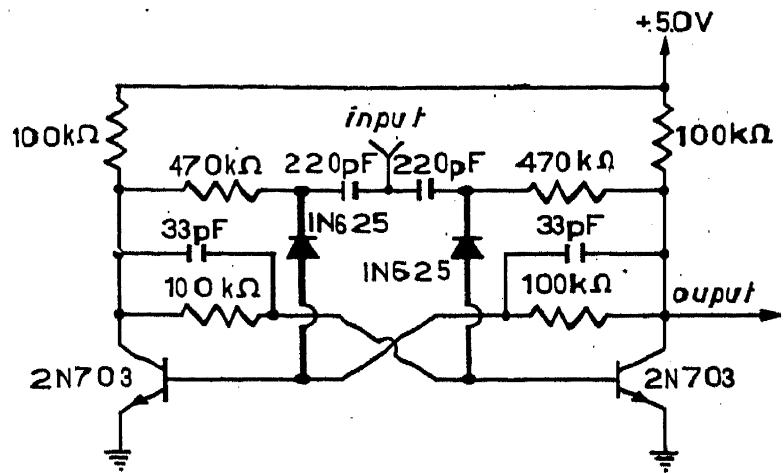


Fig. 4. — Simple 6 kHz binary scalar stage (Explorer I, III, IV, VII, S-46, Pioneer IV, V).

The counting, or scaling, circuit shown in Fig. 4 meets the requirements of simplicity and operates over a wide temperature range [2-5]. The circuit requires less than one milliwatt of power per stage. It is a very simple circuit which can be used whenever the counting rate does not exceed 5000 s^{-1} and whenever a low output impedance is not required. It is capable of driving two similar circuits with no interposed circuitry, as are the other scalars described below. Its attractiveness lies in the relatively small number of components required. It operates reliably over the temperature range -50 to $+90^\circ\text{C}$ and over the supply voltage range from 4.5 to 8 V. When higher speed operation is necessary, a modification of this circuit, wherein the collector and cross-coupling resistors are halved in value and the pulse steering circuit time constant

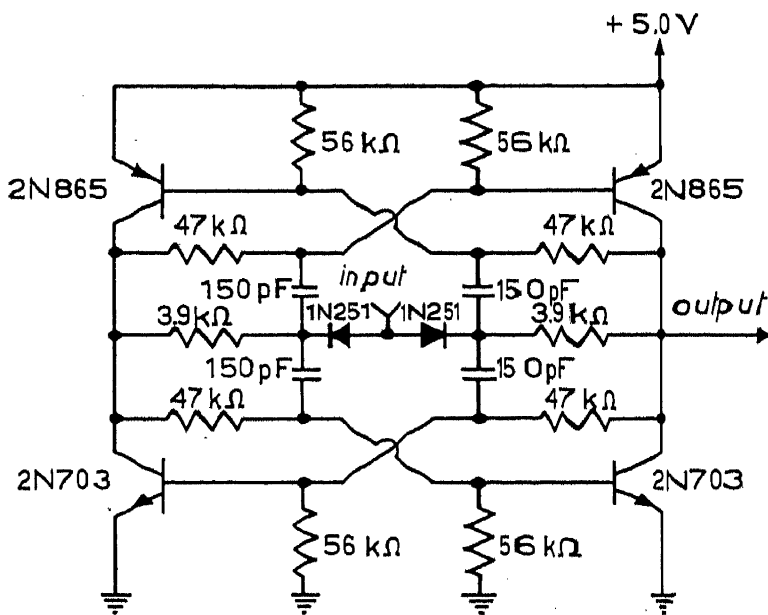


Fig. 5. — Complementary symmetrical 250 kHz binary scalar stage (Atlas-Able V, S-3).

operation is necessary, a modification of this circuit, wherein the collector and cross-coupling resistors are halved in value and the pulse steering circuit time constant

is halved, can often be used. This circuit also requires less than one milliwatt of power per stage. When an even higher operating speed or a low output impedance is required, the complementary symmetrical scaler shown in Fig. 5

can be used [6]. In this circuit, the fixed collector resistors of Fig. 4 are replaced by transistors in a configuration in which one transistor is cut off and the other is conducting in each branch. Therefore, the output impedance and the current is low in both stable states. Circuits of this type have been used at rates up to 250 kHz over the temperature and voltage ranges of -30 to $+90^\circ\text{C}$ and 4.5 to 12 V respectively, with a power requirement of about one milliwatt per stage.

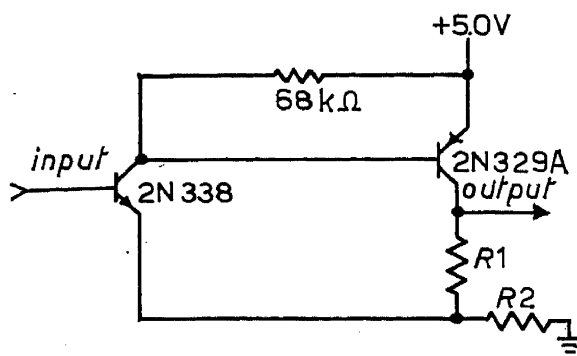


Fig. 6. - Basic bootstrap amplifier.

Linear pulse amplifiers are required to amplify by known and constant factors the pulses from ion chambers, proportional counters, scintillation detectors, Čerenkov detectors and solid state detectors. Due to the strong dependence of transistor parameters on operating conditions, it is usually necessary to employ large amounts of negative feedback to stabilize the amplifier characteristics. This also helps to improve the high frequency response of the amplifiers. Figure 6 shows a basic «Bootstrap» D.C. coupled amplifier [4], in which the gain depends almost entirely on the ratio $(R_1 + R_2)/R_2$. Figure 7 shows the gain curve for this amplifier for several temperatures. It can be seen that the AC gain, or slope, is essentially independent of temperature. Figure 8 is a schematic diagram of a pulse amplifier based on this circuit which was used in Explorer IV with a scintillation detector [7]. The thermistor in the input stage was used, not to make the amplifier characteristics independent of temperature, but to intentionally introduce a nonlinear temperature dependence to balance the photomultiplier tube characteristic to produce a flat

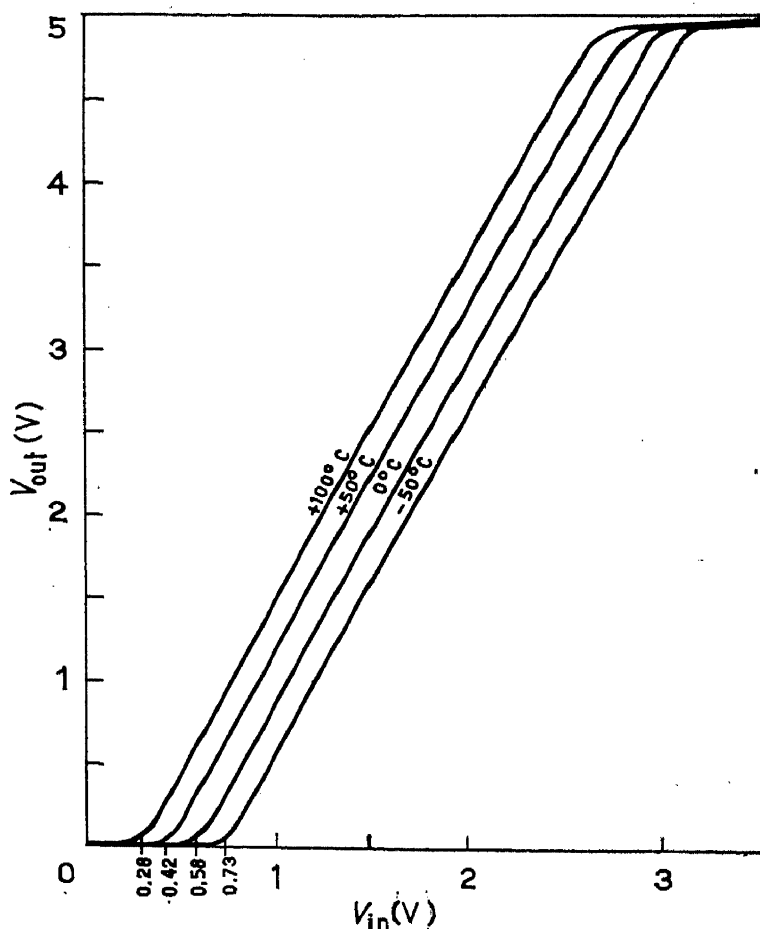


Fig. 7. - V_{out} vs. V_{in} for the amplifier of Fig. 6 with $G = 2$.

temperature characteristic for the entire system. The gain of this amplifier was 150, the rise time was $0.3 \mu\text{s}$ and its operating power was only 10 mW.

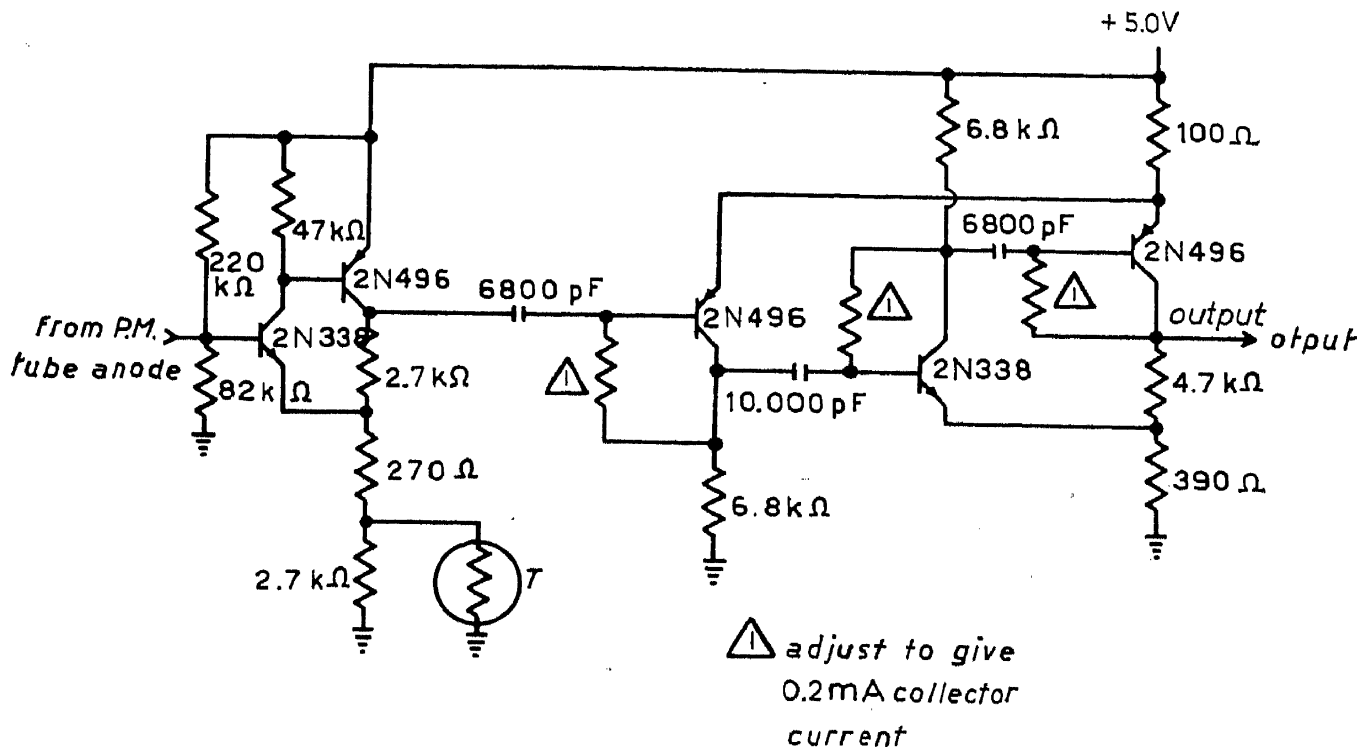


Fig. 8. — Linear pulse amplifier (Explorer IV).

Another fast linear pulse amplifier is shown in basic form in Fig. 9. This is an operational amplifier in which the gain is, to a first approximation, $-R_1/R_2$ if the first stage input impedance is high, the second stage gain is high, and the load impedance is high. A two-stage amplifier and emitter follower having an overall gain of 100 is shown in Fig. 10 (*). Both the bootstrap and operational amplifier circuits shown possess high input and low output impedances and linear operation at frequencies higher than 1 MHz. They can be used with all pulse detectors by properly adjusting the number of stages and stage gains, and by providing suitable input circuits.

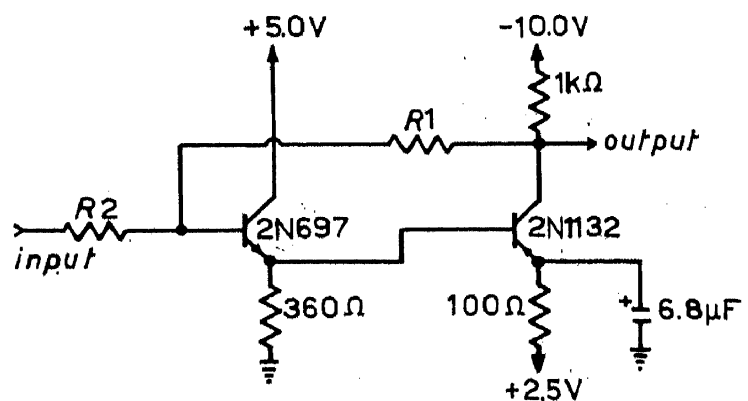


Fig. 9. — Basic operational linear amplifier.

Frequently it is necessary to convert an analog voltage or current to a

(*) This amplifier is part of the Radiation Instrument Development Laboratory 32 channel satellite pulse height analyzer developed under NASA contract.

used to convert current in the range 10^{-5} to 10^{-9} A to a pulse rate [8]. This technique has been applied to the measurement of current through cadmium sulfide solid state detectors, the integrated ion current in an ion chamber, and the integrated dynode current in a scintillator detector. By the proper choice of circuit components stable operation can be obtained for operation over a wide temperature range. The circuit as shown in Fig. 11 cannot be used to reliably measure currents smaller than 10^{-9} A due to the fact that a current of this magnitude can leak through the discharge tube without causing a complete ionization. A simple modification can be made to the basic circuit to permit measurement of currents lower by several order of magnitude. This is accomplished by superimposing a low amplitude periodic waveform on the discharge element. A circuit employing an auxiliary relaxation oscillator to generate this waveform is shown in Fig. 12 (*). This circuit is capable of measuring currents to about 10^{-12} A.

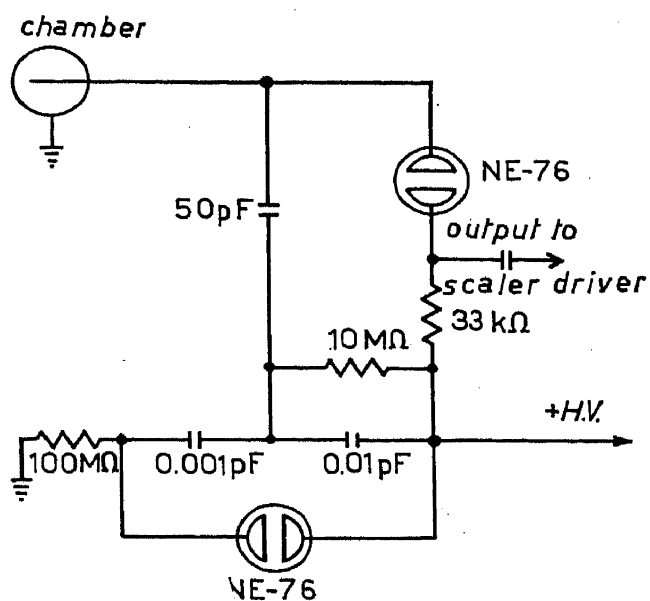


Fig. 12. - Low-current converter.

4. - Simple conditioning sub-systems.

By the use of combinations of slight variations of the basic circuits described above and in Table II, it is possible to build a large number of subsystems to perform operations at the second level of complexity. To illustrate, a voltage amplitude can be converted from analog to digital form by means of the circuit shown in the block diagram of Fig. 13. It can be used as shown for inputs having variations slow compared with the clock rate. If the input amplitude variations are more rapid, a pulse lengthener or level clamping circuit can be used to hold the level for the duration of the measurement. Action is initiated by applying a «begin conversion» pulse which opens the clock gate. The scaler counts clock pulses until the staircase level first exceeds the input amplitude. At this point a comparator output appears which closes the clock gate. The amplitude is then contained in digital form in the scalar and can be commutated, shifted out, or used directly to address a storage matrix.

(*) R. S. ROCKLIN: *General Electric Company*, private communications.

The pulse height-to-time converter shown in block diagram form in Fig. 14 contains a different type of analog to digital converter in which a capacitor

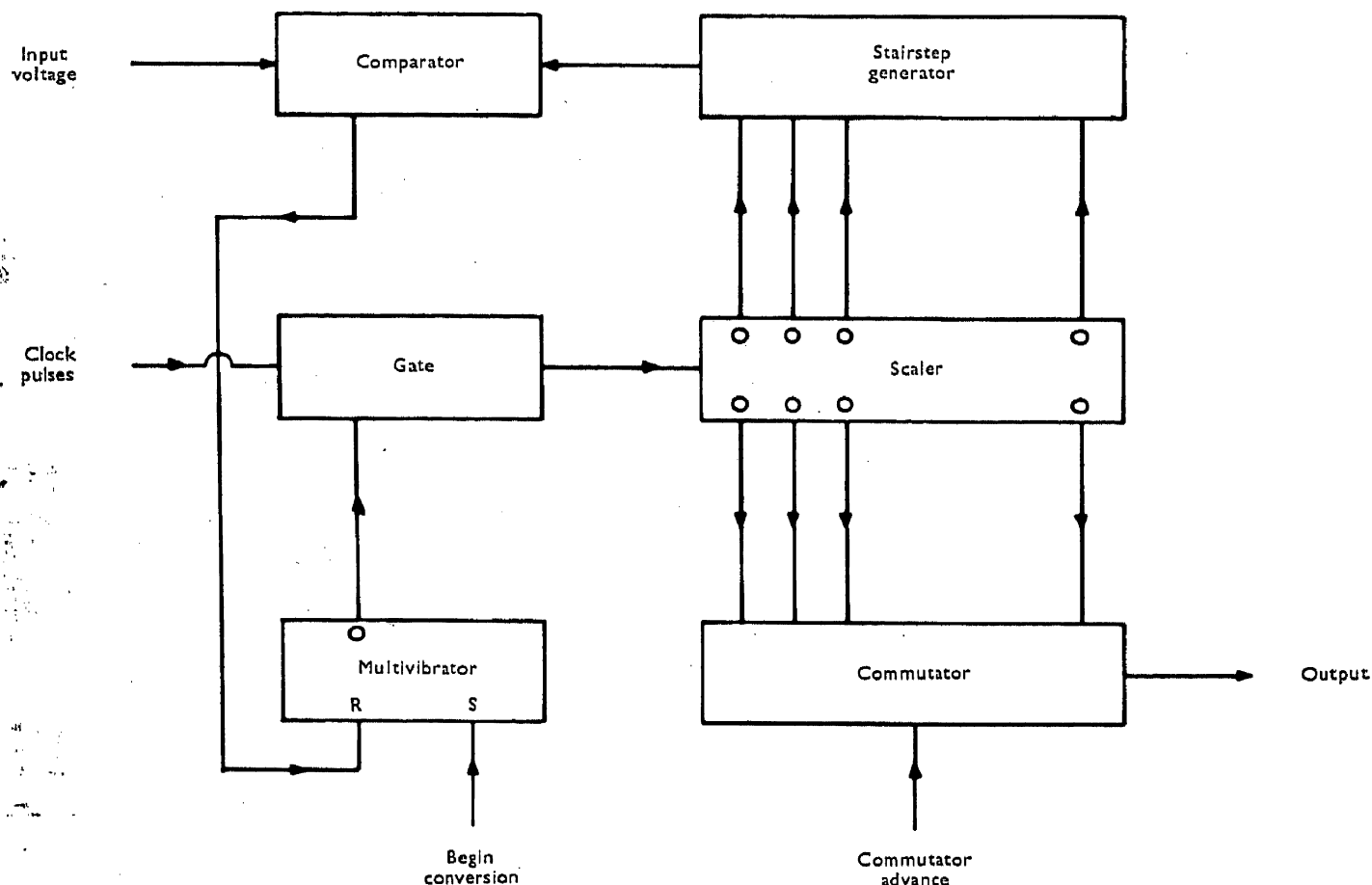


Fig. 13. - Analog voltage-to-digital converter.

is initially charged to the input amplitude and a clock gate is opened. The capacitor is permitted to discharge at a constant rate until its voltage reaches

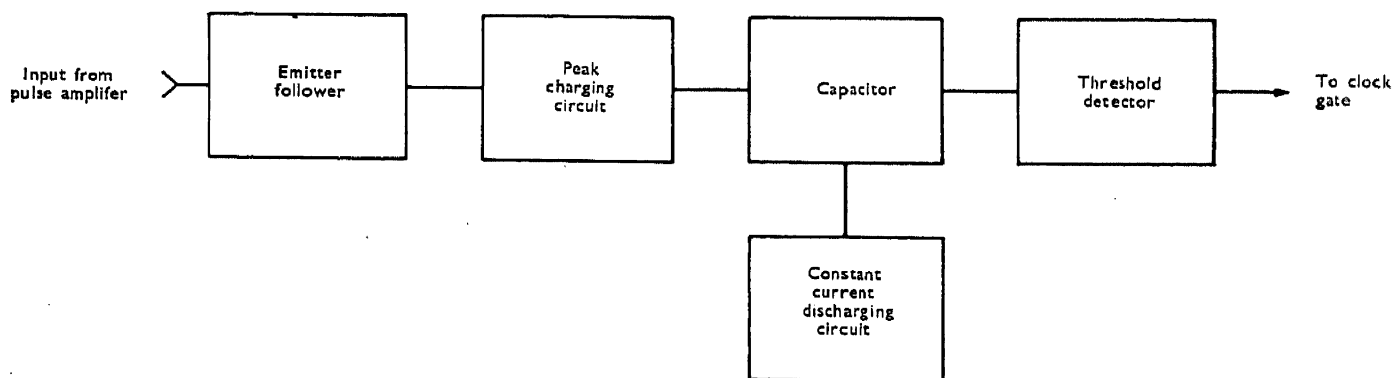


Fig. 14. - Ramp pulse height-to-time converter.

a threshold value, at which point the gate is closed. Thus, the gate width is proportional to the input voltage minus the threshold voltage. The gate allows

clock pulses to be counted by a scaler, and at the end of the conversion the digital number again resides in the scaler. The conversion speed is about the

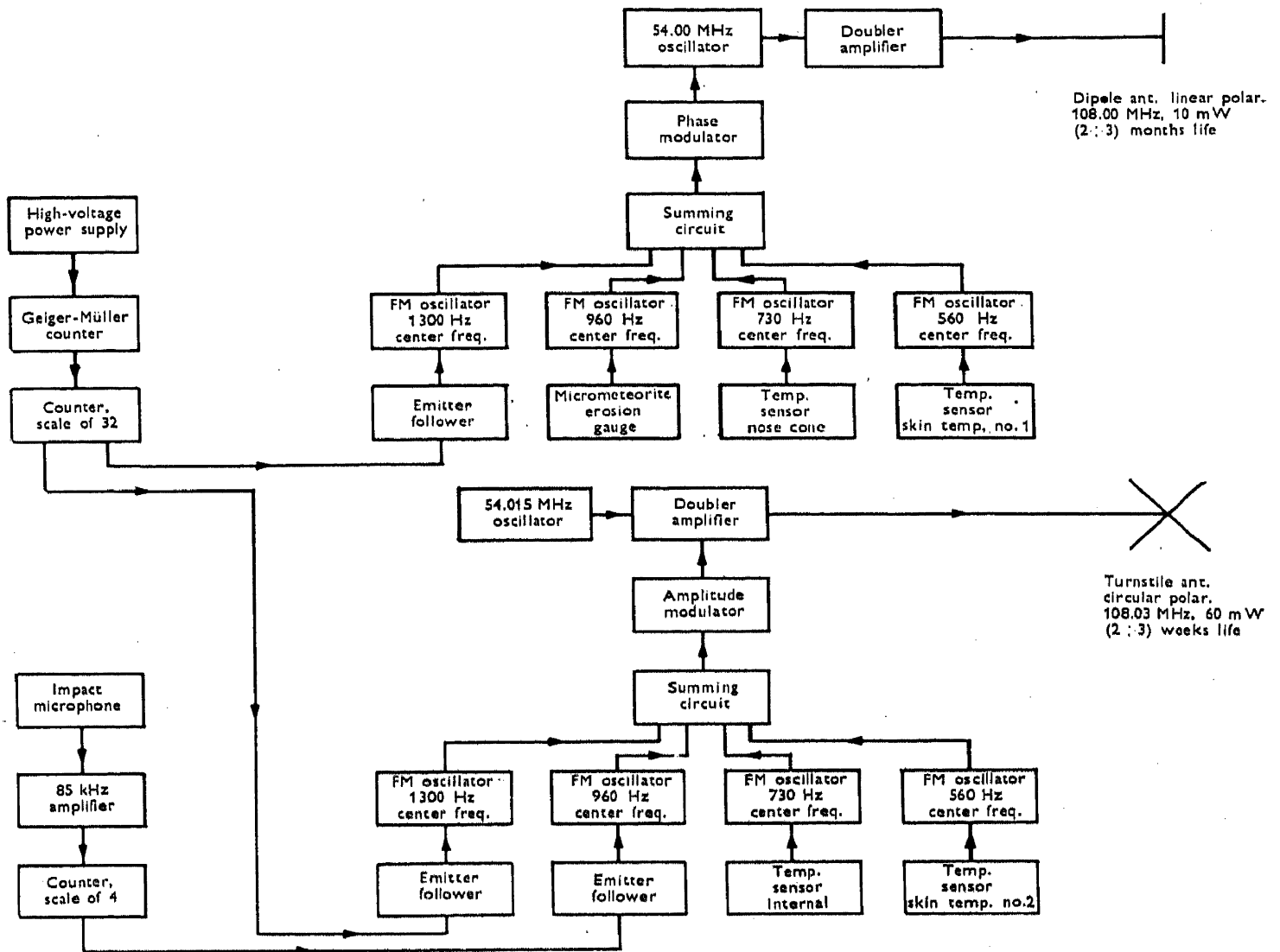


Fig. 15. — Explorer I spacecraft data system.

same as that of the circuit of Fig. 15. Both this and the previous circuit can be made with a zero offset by biasing the stair-step off-zero in the case of the previous one, and by making the threshold voltage nonzero in the latter.

5. — Complete signal conditioning systems.

In a spacecraft, the detectors and signal conditioning circuit elements must be assembled in such a manner that meaningful data can be sent to the ground receiving station for a number of detectors, often over a single telemetry system. For example, data from one detector may be accumulated and partially processed while data from another detector are being telemetered. More

specifically, the pulse height spectrum from a scintillation counter may be determined and stored in a magnetic core storage matrix while data from G.M. counter previously stored in a scaling circuit are being transmitted. Most of the energetic particle detectors can be used in similar time sharing systems. In those cases where full time telemetry is required, the data are either frequency multiplexed or a separate transmitter is utilized. To illustrate the manner in which data from a number of different detectors are processed, stored, and multiplexed onto a central telemetry link, a number of spacecraft systems will be described.

The Explorer I instrumentation is an example of an extremely simple

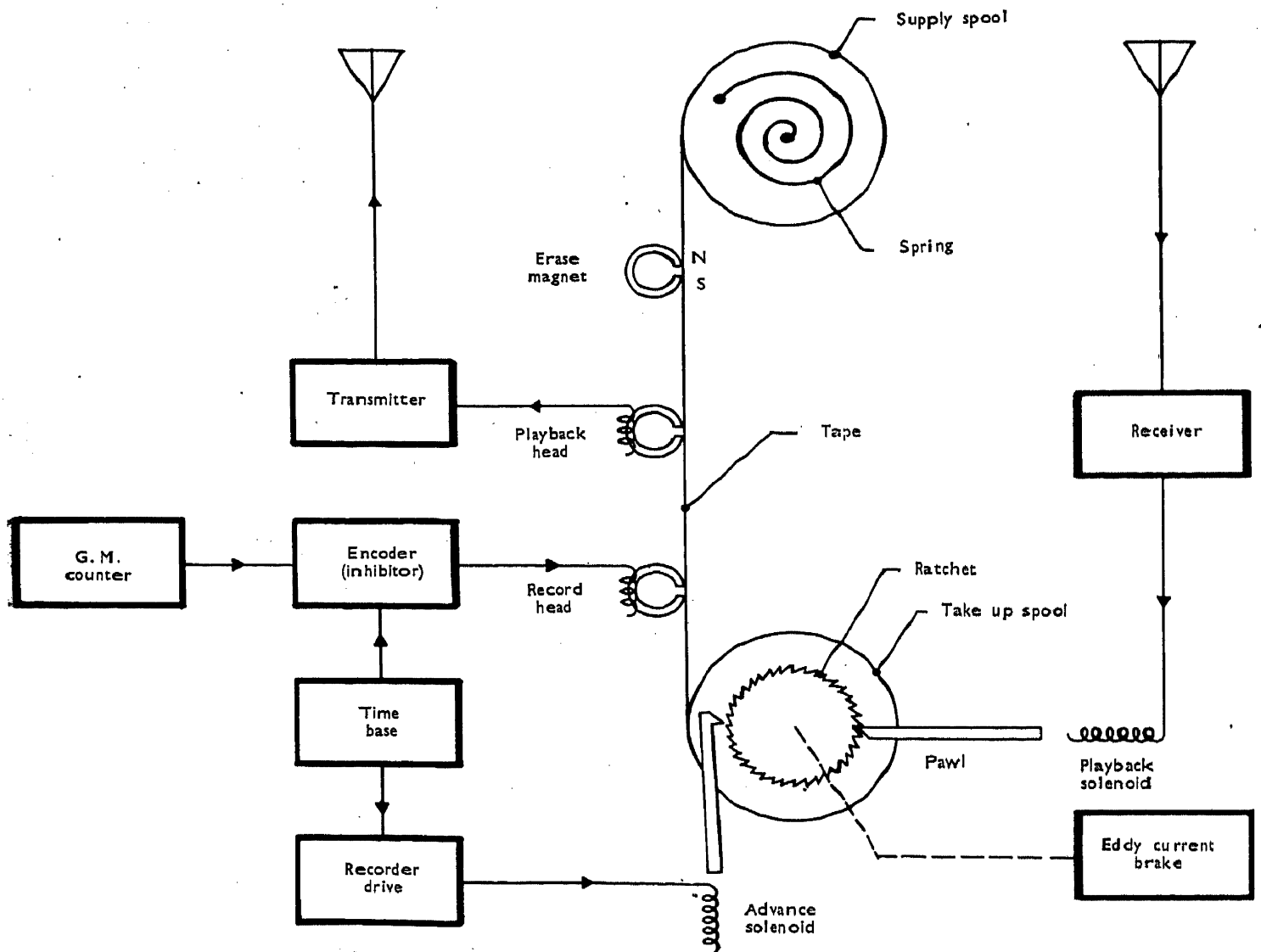


Fig. 16. - Simplified diagram of storage and readout system for Explorer III.

system in which a minimum amount of processing was provided before telemetry [2, 9]. As seen in Fig. 15, the information from each sensor directly modulated a sub-carrier oscillator. The outputs of the sub-carrier oscillators were summed in two groups, and used to modulate two independent trans-

mitters. The two independently operating systems provided redundancy to increase the reliability.

The Explorer III system is an example of a data-handling system in which very little data processing was accomplished, but in which tape recorder data storage was provided in the satellite to allow recovery of information from all positions in the orbit [10]. Figure 16 indicates schematically the method in which the recorder was used. It should be noted that the tape recorder was made digital in nature, that is to say, the tape moved in discontinuous steps, because of the essentially digital nature of the data. This resulted in a low-average power requirement for the tape recorder of only 35 mW. The data capacity was 7200 binary bits.

The instrumentation used in Explorer VI, shown in Fig. 17, is an example

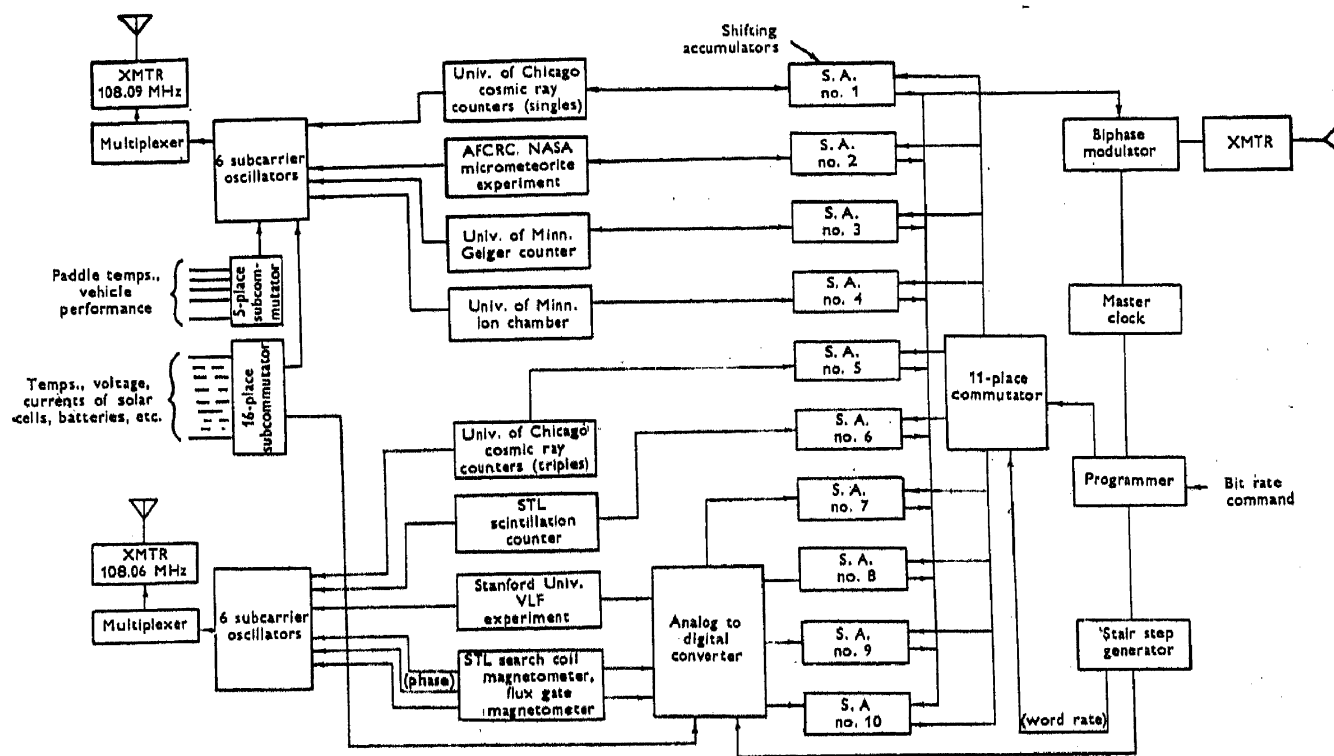


Fig. 17. - Satellite telemetry for Explorer VI.

of a somewhat more complex data system [11]. This greater complexity was a result of the fact that a larger number of experiments were carried and that a more advanced telemetry encoding system was employed. It should be noted, however, that the amount of data processing before transmission was still small. All detectors provided outputs to simple preconditioners, and the outputs of these conditioners were telemetered. The digital encoding system was a departure from previous practice. An orthogonal code and coherent modulation were employed to decrease the word error probability for low signal-to-noise ratios, as described later in this paper. Because of the fact

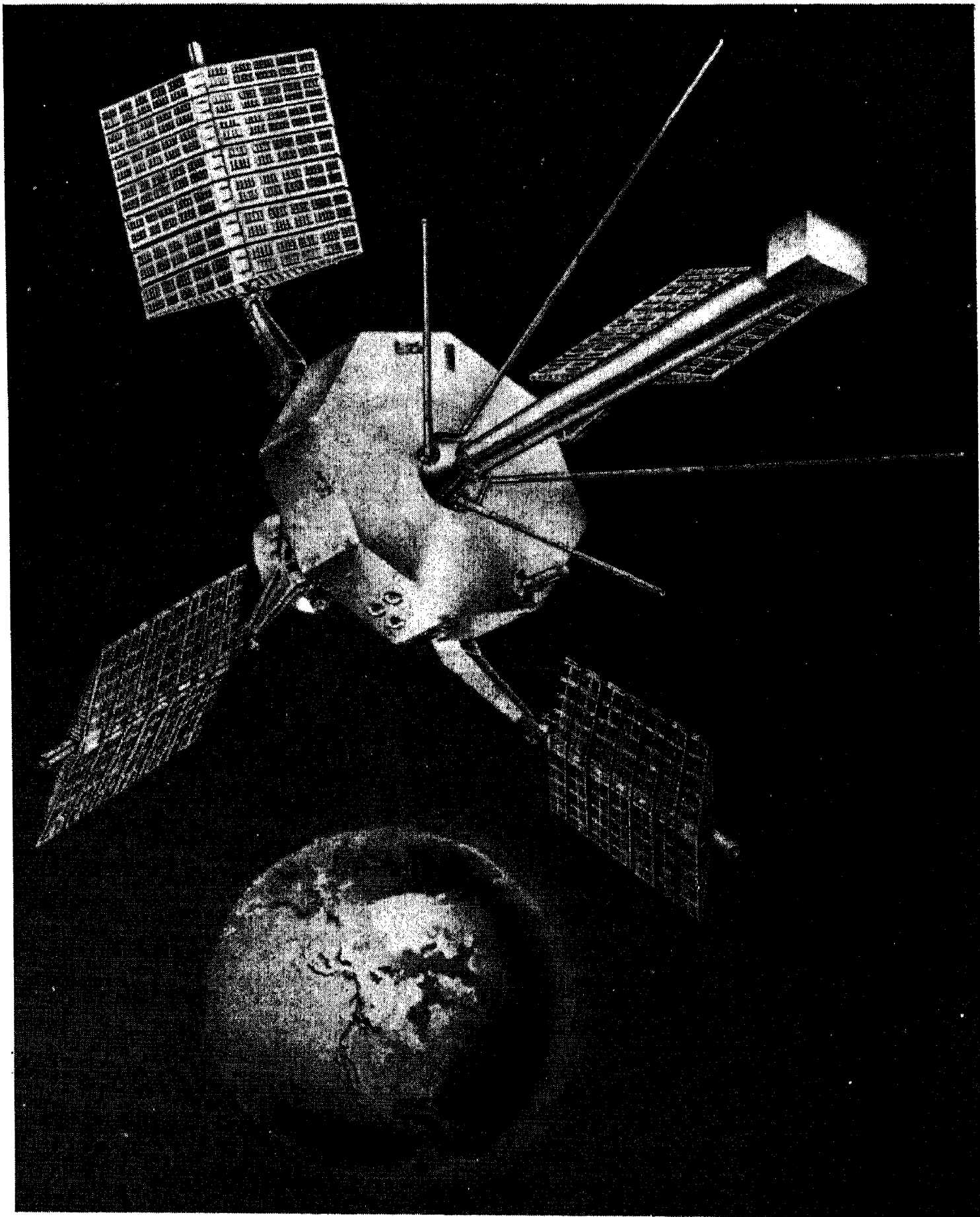


Fig. 18. - Artist view of the S-3 satellite.

that this was a new system, a separate telemetry system employing the older frequency multiplexing technique was included as a back-up system.

The S-3 satellite is an example of a present generation system in which a large number of detectors are employed and in which there is a large amount of data processing before telemetry. Figure 18 is an artists view of the complete satellite, which will be launched into a highly eccentric orbit (apogee greater than 10 earth radii) later this year. The instrumentation is designed to investigate the charged particle energy spectrum from about 100 eV to greater than 10^9 eV. Figure 19 indicates the particle flux-energy values which

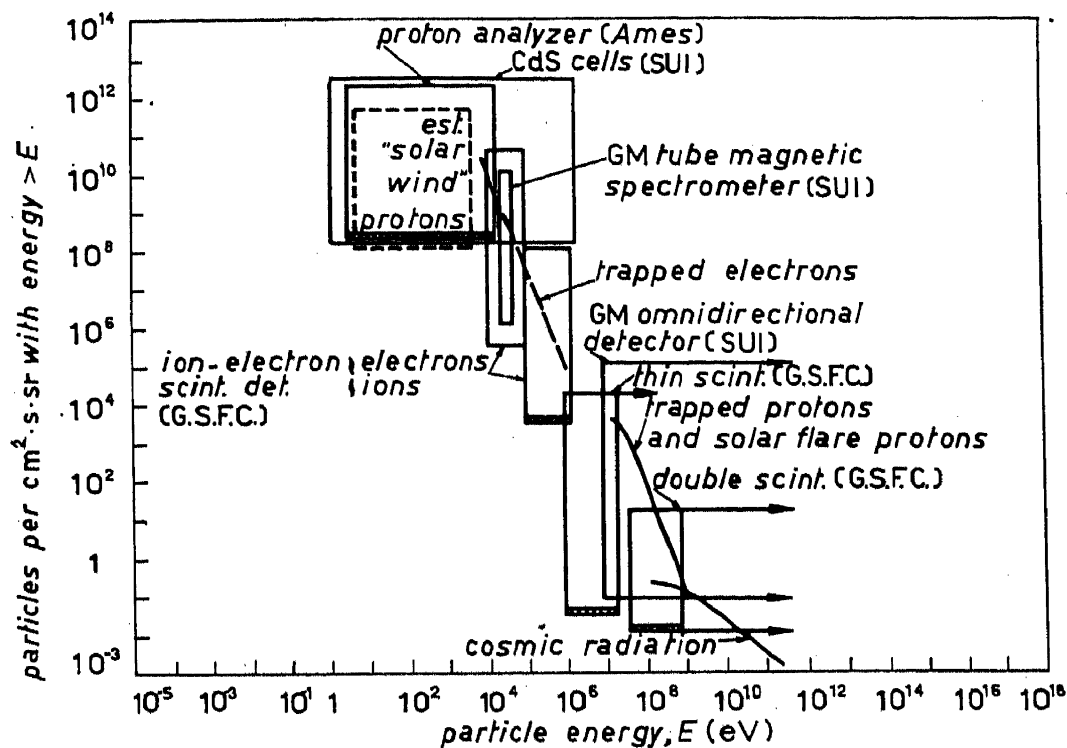


Fig. 19. - Range of sensitivity for energetic particles for instruments on satellite S-3.

will be measured. The experiments for these investigations are being furnished by the Goddard Space Flight Center, the State University of Iowa, and the Ames Research Center and a two axis flux gate magnetometer is being provided by the University of New Hampshire. The satellite structure, power system, telemetry system, thermal control system, de-spin system, technological experiments and optical aspect sensor have been developed by the Goddard Space Flight Center.

Three of the Goddard Space Flight Center experiments form a major experiment sub-system which will be described to illustrate the more complex data processing techniques employed. These experiments were designed to assist in the study of the cosmic ray accelerating mechanisms and the nature of the modulation mechanisms which result in the eleven year variation and the Forbush decreases which accompany certain types of solar activity, and to

increase our understanding of the mechanisms by which solar cosmic rays are produced and modified by solar weather. This instrumentation is shown in block diagram form in Fig. 20. The first detector is a double scintillator tele-

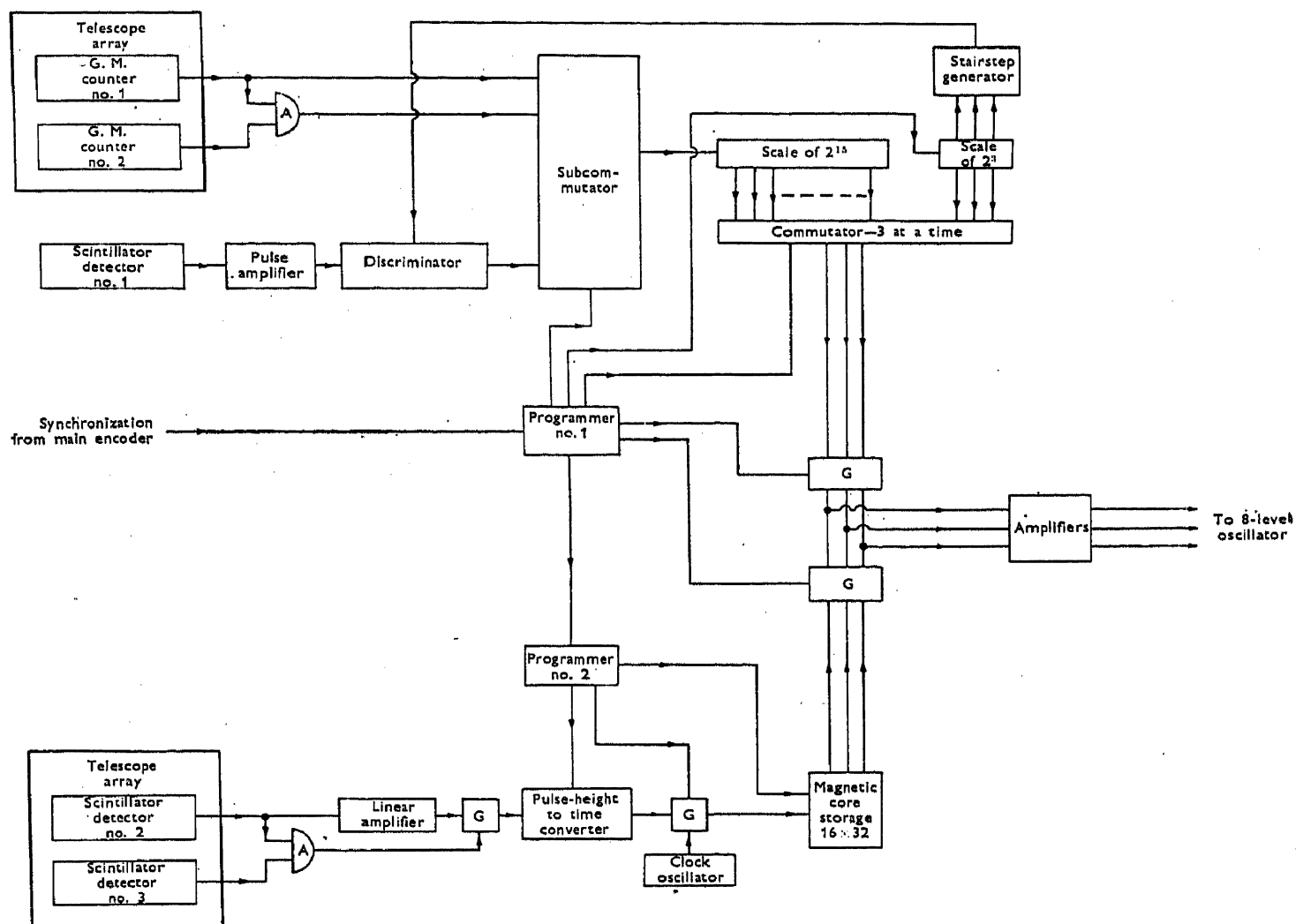


Fig. 20. — The S-3 cosmic-ray signal conditioning systems.

scope which is capable of measuring the proton spectrum in the energy range 70 to 700 MeV. Only particles traversing both elements of the assembly are processed. When a coincidence is obtained, the pulse from one of the scintillator detectors is analysed by a pulse height analyser, in which the pulses are sorted into one of 32 storage channels depending on which of 32 amplitude increments they fall within. The analyser utilizes the linear pulse amplifier of Fig. 10 and a pulse height-to-time converter of the type shown in Fig. 14. The gated clock pulses are counted in the address register of a magnetic core storage matrix. At the completion of this counting operation, one count is added to the number previously stored in the channel addressed by the address register. At the completion of the store mode, the complete pulse height spectrum (differential energy spectrum) is contained within the 16 by 32 core matrix. During the readout mode, these 512 binary bits are sequentially tele-

metered. The pulse height analyser as described above has been designed for use in spacecraft by the Radiation Instrument Development Laboratory and is available in a package measuring $(15.2 \times 15.2 \times 12.7)$ cm, weighing 2.5 kg, and requiring about one watt of power for its operation.

The second detector used in this set of instrumentation to extend the proton energy spectrum down to about 1 MeV is a thin CsI scintillator detector. The pulse height distribution from this detector is determined by an eight channel integral pulse height analyser developed at GSFC [6]. In this analyser, a pulse amplitude discriminator level is established by an eight level staircase generator. All pulses having an amplitude greater than the discrimination level are counted by a scaler. At each level the pulses are accumulated and then read out. The discriminator is sequentially stepped through its eight levels.

The third detector employs two Anton type 1003 pancake G.M. counters arranged to serve as a simple, reliable, cosmic ray monitor and to check on the information received from the other two detectors. The counters form a telescope array, and the coincidence rate and the rate from a single G.M. counter are accumulated and telemetered sequentially.

A commutating system is included to allow all three detectors to utilize common telemetry channels. At the beginning of the commutator cycle the G.M. counter coincidence events are accumulated for about 1.6 s and read into the telemetry system during the following 0.96 s. Then the single G.M. counter output pulses are accumulated and read for comparable time intervals. The single scintillator detector output pulses, with the discriminator set at the first stair-step level, are accumulated and read during the next 1.6 and 0.96 s periods respectively. This is repeated for stair-step levels two through eight. Thus, 25.6 s are required for reading the counting rates of these two detectors. This sequence is repeated 12 times, requiring a total of slightly more than 5 minutes.

During the time the G.M. counter and single scintillator detector pulses are being accumulated and read, pulses from the double scintillator telescope are being analysed and stored in the magnetic core memory. At the completion of this sequence, the data lines to the telemetry system are switched to the analyser storage system, and its readout begins.

The stored data are telemetered several times during the ensuing 102 s. Thus, a new energy spectrum and a complete set of G.M. telescope and signal scintillator detector data are obtained every 6.7 minutes.

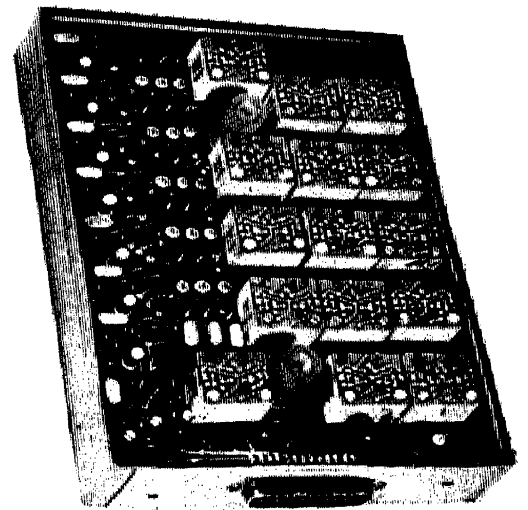


Fig. 21. — The S-3 cosmic-ray logic system accumulator card. This sub-assembly is 17.8 cm long.

The sub-system outlined above, including the pulse height analyser, was fabricated in the form of five sub-assemblies utilizing approximately 530 transistors, weighing 5.8 kg, and requiring 1.4 W of electrical power. An example of the method of fabrication is shown in Fig. 21. This circuit board contains the scaler (of the type shown in Fig. 5) used to accumulate the G.M. tube and single scintillation detector pulses, and the matrix required to read it into telemetry system.

6. - Advanced spacecraft central data handling systems.

Several data handling systems have been discussed in which the signal conditioning, storage, and data handling functions were combined into integral systems. In line with the development of the larger standard observatory spacecraft such as the Orbiting Geophysical Observatory (OGO) and Mariner, large centralized data handling systems are being developed which can accept the standardized outputs of a large number of sensors and directly associated signal conditioning systems. The OGO data system shown in Fig. 22 is an example of this type of advanced system for use in earth satellites where the transmission path length is relatively small (less than 20 earth radii). This system actually contains three sub-systems, a wide-band digital data sub-system, a special purpose analog sub-system, and a radio command sub-system. Since the digital sub-system includes analog-to-digital converters, it can accept a wide variety of digital and analog inputs. The wide-band sub-system will be capable of telemetering at one of three binary bit rates. The selection of the 1000, 8000, or 64 000 bit s⁻¹ rates can be made by ground command. The digital information will be presented in 9 bit groupings, or words, so that about 115, 920, or 7360 words, or data samples, per second can be telemetered. A commutator will be included to allow the sampling of as many as 544 inputs according to a predetermined format.

The wide-band system also includes a tape recorder storage system capable of storing nearly 85 million binary bits. Continuous storage at a rate of 1000 bits per second is being provided and read-out of the recorder at a high rate by ground command will be possible. Thus, data will be recorded from all positions in the satellite orbit.

The special purpose telemetry is being provided for wide-band experiments whose data cannot be digitized easily. The transmitter and antenna system are being provided as a part of the spacecraft, but all signal-conditioning instrumentation will probably be provided by the participating experimenters. The system will have a 200 kHz bandwidth.

The digital command system will be included to permit partial control and fault correction in the spacecraft. Commands will be provided: 1) to per-

mit overriding stabilization, thermal, and other sub-systems, 2) to permit control of the data system, 3) to turn experiments on and off, and 4) to control scaling factors, ranges of operation, absorbers, etc., in the experiments. A total of about 250 commands will be possible.

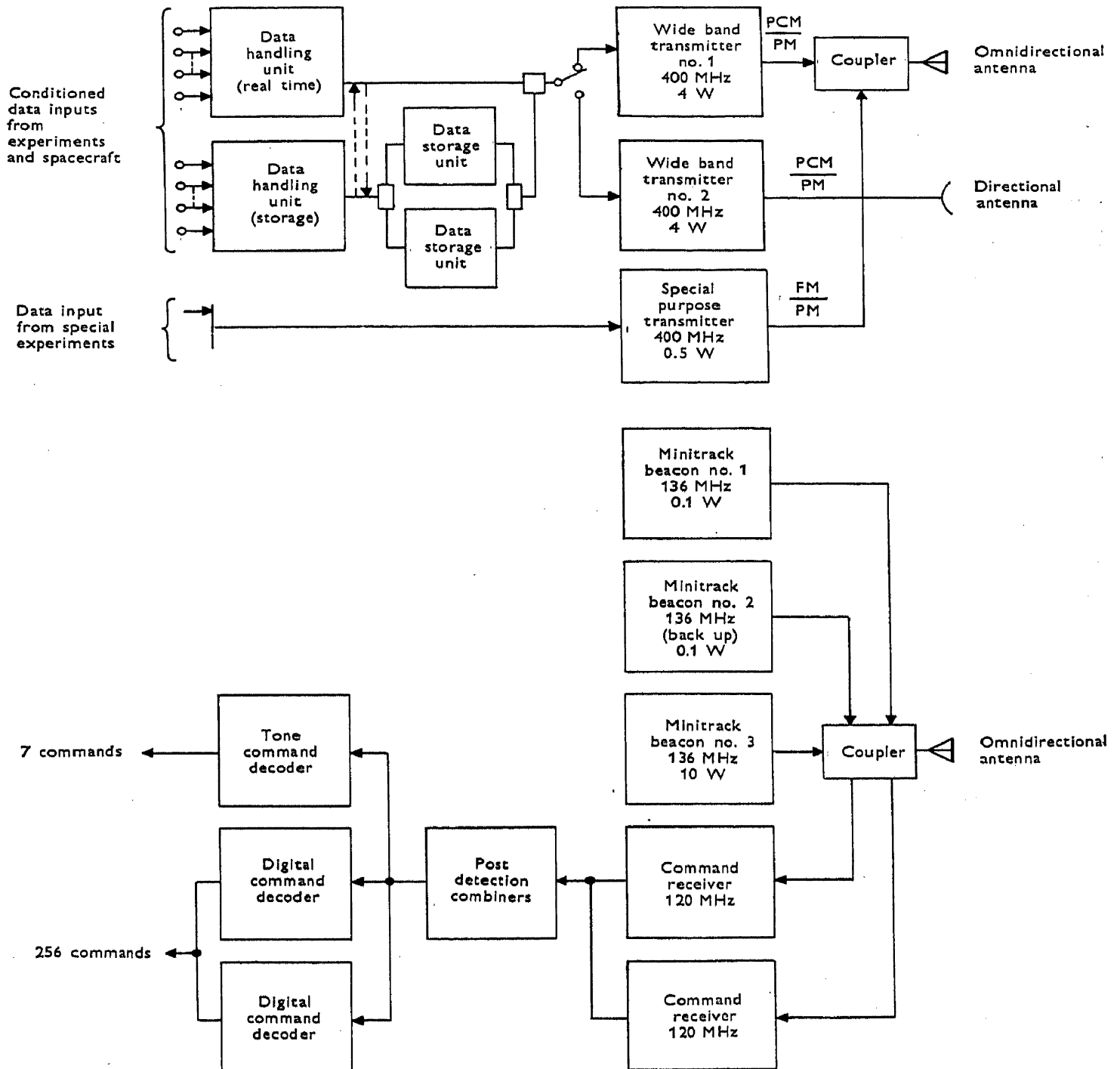


Fig. 22. - The Orbiting Geophysical Observatory data system.

In the development of these central spacecraft data-handling systems, the goal is to make them as universal as possible and to make the interface between the experimental instrumentation and the data-handling systems as

simple and well defined as possible. This is expected to reduce the effort required by the experimenters in making instrumentation for a large number of different spacecraft and in making equipment which will work with the data systems properly the first time they are mated.

7. - Telemetry.

Once the detector signals are processed into a form suitable for transmission, the next problem is that of telemetry. The goal is to obtain as much power per data bit at the receiver output per unit transmitter input power as possible, while keeping the noise at the receiver output as low as possible. The familiar equation for signal-to-noise power ratio is developed as follows [12]: The power area density at a point in space (magnitude of the Poynting vector) is:

$$P = \frac{W_t G_t}{4\pi R^2}$$

and the power intercepted by a receiving antenna is:

$$W_r = \frac{P A_r}{l}$$

Thus, the received power is:

$$W_r = \frac{W_t G_t A_r}{l 4\pi R^2}$$

where W_t is the transmitted power,

G_t is the gain of the transmitting antenna,

A_r is the *effective* area of the receiving antenna,

R is the distance between transmitting and receiving antennas, and

l includes the losses due to polarization misalignment, cabling, etc.

Note that the received power is stated in terms of the *gain* of the transmitting antenna and the *effective area* of the receiving antenna. Either the gain or area of either antenna could be used, since the gain and effective area of any antenna are related by the expression

$$A_r = \frac{G_r \lambda^2}{4\pi},$$

where λ is the wavelength. Table III lists the gains and effective areas of several common antennas. The question of whether A_r or G should be used

depends only on which is the more significant for a given antenna. An isotropic antenna is defined as having unity gain (for all frequencies) and is therefore most easily characterized by G . A parabolic antenna, on the other hand, is more conveniently characterized by its effective area, since A_e is simply related to its geometric area, while its gain is a function of frequency.

TABLE III. - *Power gains and effective areas of several common antennas.*

Radiator	Gain above isotropic	Effective area
Isotropic	1	$\lambda^2/4\pi$
Infinitesimal dipole or loop	1.5	$1.5(\lambda^2/4\pi)$
Half-wave dipole	1.64	$1.64(\lambda^2/4\pi)$
Parabola (geometric area = A)	$(6.3 \text{ to } 7.5)(A/\lambda^2)$	$(0.5 \text{ to } 0.6)A$
Broadside array (area = A)	$4\pi(A/\lambda^2)(\text{max})$	$A (\text{max})$
Turnstile	1.15	$1.15(\lambda^2/4\pi)$

The selection of an antenna for the ground station depends on a number of factors [13, 14]. For reception from satellites while they are within several hundred kilometers from the earth, where the initial acquisition must be rapid and the maximum angular rate at the receiving station may exceed two degrees per second at the zenith, and for presently used frequencies, the antenna size is limited by our ability to move it rapidly. Present antennas for this purpose located at the stations which make up the Minitrack net are made up of stacked Yagi antennas with total power gains of about 20 db. For spacecraft at distances of more than several earth radii, the limiting factor is a different one. The parabolic antenna is presently capable of giving larger gain for this application than any other type for a given expenditure of effort. But, for the relationship between the effective and geometric areas listed in Table III to hold, the mechanical surface irregularities of the reflector must not exceed about $\lambda/16$. Thus, as the parabolic antenna is made larger and larger (with constant frequency) a point is reached where a further increase in size will not result in a proportionate increase in antenna gain unless a large refinement in the antenna construction tolerance is made. Figure 23 shows the relationships between the antenna power gain and frequency for two parabolic antennas in present use, the 250 foot diameter Jodrell Bank antenna and the 85 foot diameter Goldstone, Calif. antenna. The ratio of surface irregularity σ to diameter d of 10^{-4} for the upper curves represents a construction tolerance of about $\frac{3}{4}$ cm for the 250 foot antenna and about $\frac{1}{4}$ cm for the 85 foot dish, while the value $2.5 \cdot 10^{-4}$ represents tolerances of about 2 and 0.6 cm respectively.

On the spacecraft the factors limiting antenna size and gain are considerably different. If the spacecraft antenna is not oriented toward the earth then a radiation pattern approaching isotropic is necessary. Even if the space-

craft is stabilized, it will not be possible for a number of years to make its antenna as large, as accurate, and as controllable as the earthbound antenna. During the next few years, areas of tens of square meters may be possible, and later, areas of several thousand square meters may be achieved. But it is expected that the size will always be small compared with that of the earth-bound antenna.

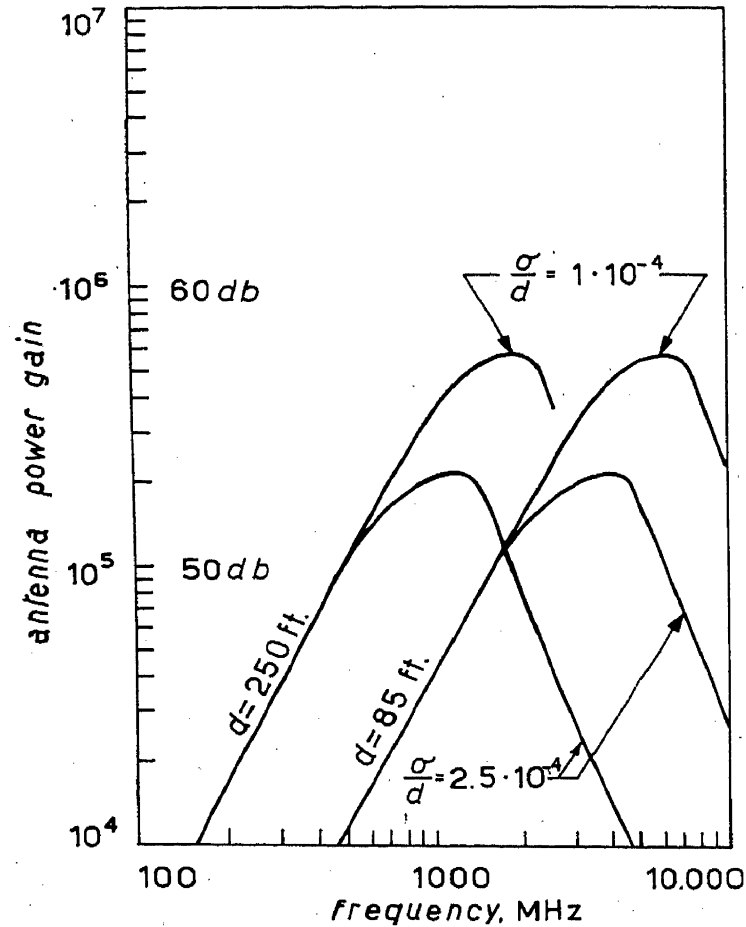


Fig. 23. — Antenna power gains vs. frequency for 85 and 250 foot diameter parabolic antennas. The antenna diameter is d , and σ is the size of the surface irregularities.

The factor l in the expression for the received power represents the losses due to the cabling from the ground antenna to the receiver pre-amplifier, and to misalignment of the planes of polarization of the incident wave and the receiving antenna. The cabling losses can be expected to be several decibels (db) while the loss due to polarization misalignment depends on the type of polarization employed in the two antennas and on their relative alignment. If the spacecraft is stabilized and if the polarization plane of the ground antenna can be rotated to correspond with that of the incident wave, then the polarization loss can

be virtually eliminated. The power attenuation factor caused by polarization vector misalignment in the general case is [12]:

$$\frac{W}{W_{\max}} = \frac{(1 \pm r_1 r_2)^2 + (r_1 \pm r_2)^2 + (1 - r_1^2)(1 - r_2^2) \cos 2\theta}{2(1 + r_1^2)(1 + r_2^2)}$$

where $W_{\max}$ is the power received when the vectors are aligned and rotating in the same direction,

r_1 is the axial ratio of the elliptically polarized incident wave,

r_2 is the axial ratio of the elliptically polarized receiving antenna

θ is the angle between the directions of maximum amplitude of the incident wave and antenna, and

— is to be used if the rotational sense is the same for the incident wave and antenna while the minus sign is to be used if they are opposite.

Several degenerate cases for this expression are summarized in Table IV.

It should be noted that the polarization of the incident wave was discussed. This will, in general, not be the same as that of the transmitted wave due to atmospheric effects.

TABLE IV. — *Polarization losses for linear and circular polarization.*

Polarization of incident wave	Polarization of receiving antenna	Sense of rotation	$W/W_{\max}$
Circular	Circular	Similar	1
Circular	Circular	Opposite	0
Circular	Linear	—	$\frac{1}{2}$
Linear	Circular	—	$\frac{1}{2}$
Linear	Linear	—	$\cos^2 \theta$

The parameter W_t is the power actually radiated by the transmitting antenna, and is less than the total input power to the transmitter by the amount of the cabling losses and the losses due to transmitter inefficiency. At the present time, it is possible to achieve transmitting system efficiencies between 30 and 80% for transmitters below 1000 MHz, and efficiencies on the order of 10 to 50% for transmitters above 1000 MHz. These figures are not likely to be improved appreciably. In computing the received *information* power it is necessary that the *usable sideband* power rather than the total radiated power be used. But in computing the received power for *tracking* purposes where only the carrier power is utilized, then only the radiated carrier power must be used. If the same transmitting system is to be used for both tracking and the transmission of data, then a ratio of carrier to usable sideband power should be selected which will result in comparable reception thresholds for the two signals. This ratio can be optimized by a proper choice of modulation method and modulation index.

If there were no interference, it would be possible to receive the signal from any transmitter from any distance by incorporating sufficient amplification in the receiver. Interference is, however, always present and it limits the usable range of communication [13-15]. It enters at every point in the telemetry link. In the transmitter, it is caused by the instability of the oscillator. In the transmitter-receiver space link, it enters from galactic, solar, planetary, and atmospheric sources. And in the circuits of the receiver, interference is generated by the random motion of the electrons and by oscillator instabilities of the type that occur in the transmitter. Every effort is made to reduce the effect of these interfering sources in order to produce as high a ratio of received signal power to total interference power as possible. For

most spacecraft systems, the interference produced by oscillators, both in the transmitter and receiver, can be made much smaller than that from other sources. The remaining interference is of two types, external and thermal. The external noise includes mostly man-made noise which is reduced by locating the receiving stations as far from civilization as possible, by careful design of the antenna, and by proper choice of distinctive coding methods to allow reading the data through the noise.

The thermal noise imposes a fundamental limitation due primarily to the random motion of electrons. The noise power produced by a thermal source is

$$W_{NT} = kTB,$$

where W_{NT} is the noise power,

k is the Boltzman's constant $= 1.38 \cdot 10^{-23}$ W s deg $^{-1}$,

T is the absolute temperature in $^{\circ}\text{K}$, and

B is the bandwidth in Hz.

It is obvious that there are two ways of minimizing the noise power, by reducing either T or B . Since reduction of B implies reduction of the telemetered information rate, every attempt should be made to make T as low as possible. The temperature T is the temperature of the equivalent resistor which represents a noise source. It also has a significant relationship to the equivalent temperature of those portions of earth, atmosphere, planets, sun, and galaxy which are seen by the receiving antenna. The value of T to be used in the equation is the properly weighted sum of all the source temperatures, where the weighting depends on how the source is «seen» by the receiving system. For sources which are internal to the receiving system or uniformly surround the antenna, the weighting factor is one. For sources which occupy only a fraction of the antenna pattern, the weighting factor depends on the size of the source and where within the antenna pattern it is located. The weighted temperature T_{nw} for a source n is given by

$$T_{nw} = \frac{1}{4\pi} \iint T_n(\varphi, \theta) G(\varphi, \theta) d\varphi d\theta$$

where T_n is the equivalent temperature of source n ,

G is the antenna gain, and

φ and θ are angular co-ordinates, measured with respect to the central antenna beam axis.

The temperature T , then, is simply $T = \sum_n T_{nw}$.

Typical values of T_n for the various sources are as follows:

1) Receiver noise. The equivalent noise temperature for conventional, parametric-amplifier, and maser receivers are 2000 °K, 100 °K and 10 °K respectively.

2) Earth black-body radiation. If the antenna «sees» the earth within a portion of its radiation pattern, then noise due to the black-body radiation of the earth at a temperature of about 280 °K is received. This source is essentially frequency-independent. Antennas whose sidelobe levels are low and which accept very little of this noise are called low-temperature antennas. They are characterized by equivalent temperatures of from 10 to 20 °K. A more conventional well-made parabola and feed may have a temperature of about 100 °K.

3) Atmospheric noise. The atmospheric noise is a strong function of frequency and of path length through the atmosphere, as shown in Fig. 24. For angles above the horizon greater than 10°, the equivalent noise temperature is less than about 10 °K. It can be seen that for good maser systems the performance will be seriously degraded below 10° elevation angle because of the atmospheric noise and because the antenna sidelobes will intercept the earth. The consequences are not serious for deep space communication where the elevation angle can be kept greater than 10°, but may be for low-altitude satellites since proportionately more time is spent below this elevation angle.

4) Galactic (cosmic) noise. The galactic noise is a strong function of frequency and of the direction in which the galactic plane is viewed, as shown in Fig. 24 and 25. At 270° celestial longitude in the plane of the ecliptic, the direction is very nearly toward the galactic center, and a high temperature results from the integrated effect of the large number of discrete sources seen.

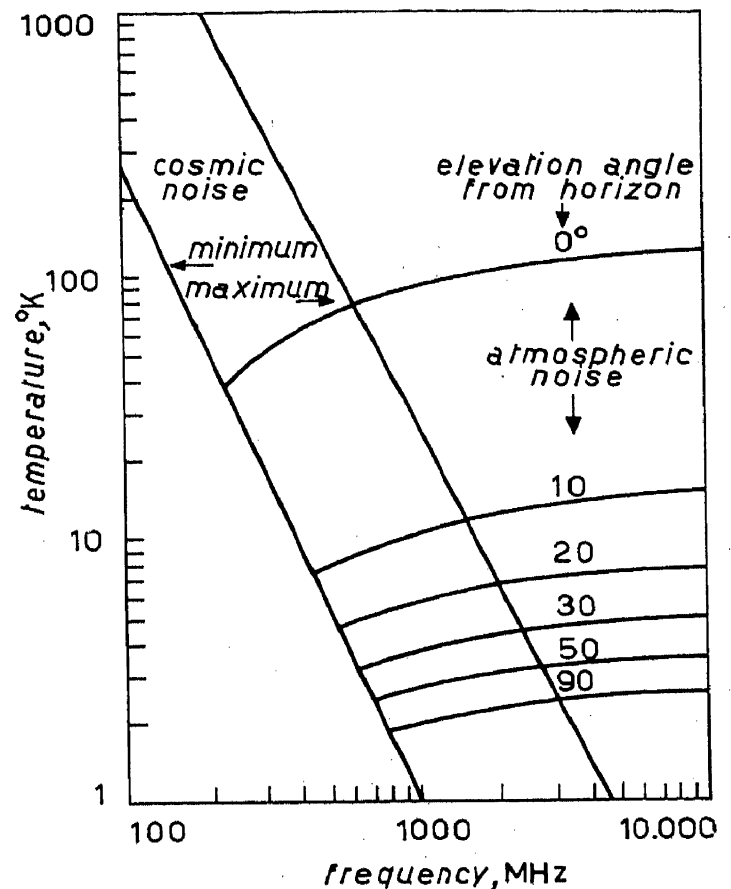


Fig. 24. Atmospheric and cosmic noise vs. frequency.

The small peak at 90° is the other interception of the galactic and ecliptic planes. For other directions, the noise temperatures for receiver frequencies of 100 and 250 MHz are about 500 and 30°K respectively.

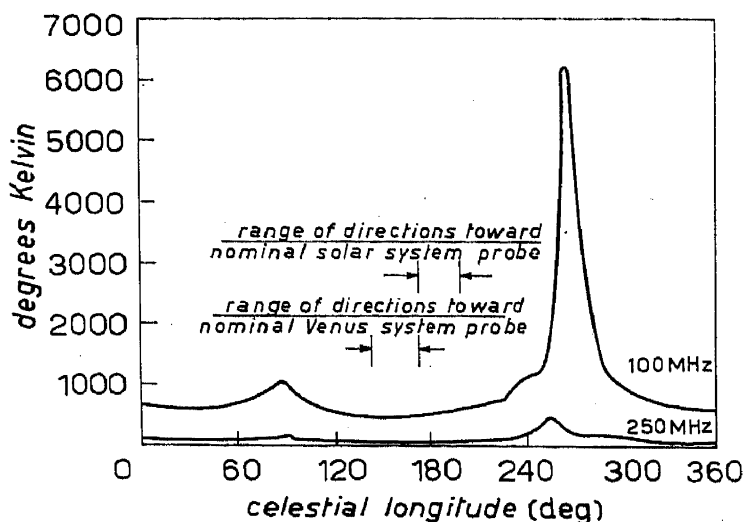


Fig. 25. — Galactic radio noise in the plane of the ecliptic *vs.* longitude. Longitude is measured east from the vernal equinox.

5) Radio star, solar, lunar, and planetary noise. In addition to the diffuse galactic source discussed above, there are a number of discrete sources, generally less than 1° in extent. Included are radio stars, which may not necessarily be identified with visible objects, the sun, the moon, and the other planets. In speaking of these sources, which occupy a very small portion of the antenna beam, it is customary to talk of their energy flux-density S at the position of the earth rather than T , where

$$T = \frac{\lambda^2 S}{8\pi k} \quad \text{and} \quad W_{ND} = SA_r B$$

where W_{ND} is the noise power from a discrete source, A_r is the effective antenna area, and B is the bandwidth, as before.

The most intense radio star known is *Cassiopea A*. At a frequency of 20 MHz, it produces a flux density of $5 \cdot 10^{-22} \text{ W(m)}^{-2} \text{ Hz}$. The flux density decreases with increasing frequency, reaching values of $2 \cdot 10^{-22}$ and $6 \cdot 10^{-24} \text{ W(m)}^{-2} \text{ Hz}$ at 100 MHz and 10 000 MHz respectively. The next most intense source produces only about half as much flux density as *Cassiopea A*, and only four of the remaining ones reach $10^{-23} \text{ W(m)}^{-2} \text{ Hz}$.

The noise power received from a radio star will be less than that received from a rather weak galactic region for antenna diameters less than about 100 feet.

The solar electromagnetic radiation is both complex and variable, as shown in Fig. 26. It consists of thermal radiation at a black-body temperature of $6\,000^\circ\text{K}$ on which additional noise is superimposed due to solar activity. Thus, it can be seen that the sun is a very intense source, and takes on secondary importance in most cases only because it is localized.

The moon and planets may be disregarded as noise sources for most practical considerations. Some of the planets are sources of noise bursts at certain frequencies which can be avoided. The thermal noise flux density for the

moon has been shown to be about $10^{-21} \text{ W(m)}^{-2} \text{ Hz}$ at a frequency of 35 000 MHz, but is presumably immeasurably small at frequencies below the microwave region. Of the planets, Venus has the highest thermal flux of $10^{-24} \text{ W (m)}^{-2} \text{ Hz}$ at 9 500 MHz. As in the case of the moon, the planetary thermal fluxes are measurable only at microwave frequencies.

The remaining parameter affecting the amount of noise power and the receiver output is the bandwidth. Both the pre-detection band-width and the post-detection band-widths must be considered. They are approximately equal to the smallest bandwidth encountered in the circuits, as long as most noise is introduced ahead of these circuits and all circuits are operated linearly. The value of B to be used in the above noise computation is approximately equal to the smaller of these figures. Obviously B should be the smallest value which will still pass the desired information. The relationship between B and H , the information bit rate, depends on the types of modulation and detection employed.

It is now possible to write the complete expression for the signal-to-noise power ratio at the output of the receiver,

$$\frac{W_r}{W_N} = \left(\frac{S}{N} \right)_r = \frac{W_t G_t A_r}{4R^2 kTB},$$

where all the parameters have been discussed above. A threshold signal-to-noise ratio is usually defined as $(S/N)_r = 1$. Usable data have been received, when S/N has been as low as several db. It is usually true that $S/N = 10$ db is considered marginal and $S/N = 15$ db is desirable when machine data reduction systems are used.

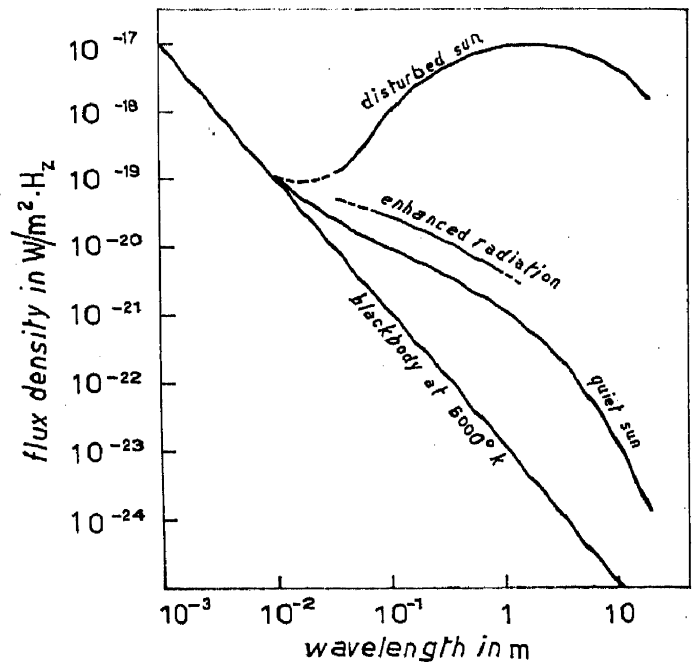


Fig. 26. — Representative solar noise levels as a function of wavelength.

TABLE V. — Representative telemetry system performance.

Spacecraft	W_t watt	Transmitter antenna gain	Receiver antenna gain	λ (m)	R (km)	B (Hz)	$(S/N)_p$
Explorer I	0.01	1.64	50	2.78	$4.8 \cdot 10^3$	20	3 db
OGO	5	16.00	10^4	0.75	$9.6 \cdot 10^4$	$2 \cdot 10^5$	15 db
Pioneer V	5	1.15	$2 \cdot 10^5$	0.31	$3.2 \cdot 10^7$	10	3 db

Table V shows the relationships between the various parameters for several sample spacecraft systems, in order to produce a better comprehension of the numbers involved.

8. - Encoding and detection.

The expression above for the signal-to-noise power ratio is true for systems using any modulation technique, as long as the values of the various parameters are the same. This is not, however, a true index of telemetry system performance, since one is really interested in the ratio of the amount of information transmitted per unit time to the total transmitter input power. Several factors in addition to the ones discussed above cause systems to give different performance. The first is the relationship between the information and receiver bandwidths. It is customary to express this in terms of the parameter $\alpha = B/H$, where B is the receiver bandwidth as before, and H is the information rate in bits per second. The bandwidth should be made just large enough to permit recovery of the information. The second factor is the ratio of the usable radiated sideband power to the total transmitter power output.

A comparison of the various encoding methods usually results in the conclusion that the most efficient one for space telemetry is pulse code modulation [16]. This is a modulation in which a pulse amplitude, phase, or frequency sequentially assumes one of a number L of discrete values. In many cases this will be simple binary code (off-on or phase reversal) modulation, in which $L=2$. Alternatively, a code can be made up of L words, each one consisting of a number M of binary bits. If $2^M > L$, then redundancy is introduced and techniques can be used which result in relatively low error probabilities for low values of S/N . The moderate complexity introduced by the use of these codes is warranted in the case of small satellites, deep-space, and planetary spacecraft on which it is difficult to achieve usable information rates by use of the more conventional techniques. A particularly interesting code, known as an orthogonal code, is one for which $L=M$, and in which the incoming word is correlated simultaneously with all L possible words [17]. There are L correlator outputs, on which appear zeros for the wrong possibilities, and A for the correct one (under noise-free conditions). A decision device selects the line having the greatest output. If noise is present so that one digital bit in the incoming word is incorrect, there will still be a high probability for selecting the proper word because of the redundancy. This decision process is known as maximum-likelihood detection, and can be shown to minimize the probability of error when all the words are equally likely and contain equal energies.

The carrier detection method must be carefully selected. It has already been stated that all receiver circuits, including the detector, should be linear, especially at low signal levels. A commonly used linear detection method is cross-correlation detection [18], in which the incoming signal plus noise ($S+N$) is multiplied by a locally generated best estimate of signal S^* . This product gives

$$S^*(S + N) = SS^* + NS^*$$

which, when averaged by a low-pass filter, yields the correlation coefficient $(SS^*)_{ave}$ contaminated by the noise $(NS^*)_{ave}$. This process is linear in that the output signal-to-noise ratio

$$\frac{SS^*}{NS^*}$$

is the same as that at the input. By comparison, in conventional square law detection, the detector output at low signal levels is

$$(S + N)^2 = S^2 + N^2 + 2SN.$$

In this process, the output signal-to-noise ratio is

$$\frac{S^2}{N^2 + 2SN} = \frac{S}{N} \frac{1}{(2 + (N/S))}.$$

Thus the degradation of signal-to-noise ratio for square-law detection compared with correlation detection is $(2 + (N/S))$. A practical method of constructing a correlation detector is by the use of a phase-locked loop.

Related to the problem of detection is one of signal coherence. A coherent system is one in which the receiver is assumed to know at all times the phase of the unmodulated transmitted signal. The *a priori* knowledge of the signal characteristic is greater for a coherent signal than for a non-coherent signal; thus, we may expect some advantage if the detection system is designed to take advantage of this knowledge. An example of a coherent system is one employing bi-phase carrier modulation, in which the carrier wave phases are 180° apart for binary 0's and 1's.

10. - Data recovery.

It is necessary to perform numerous functions at the stations where the spacecraft data are received. As shown in Fig. 27, these may include:

- 1) Data recovery. Receivers and primary data recorders for the receipt of the spacecraft data are necessary.

2) Spacecraft command. In many cases the spacecraft contains a command system, which may be used to *a)* control the transmitters, recorders and other elements of the spacecraft data system, *b)* control other spacecraft subsystems such as the stabilization system (to direct a telescope at a new

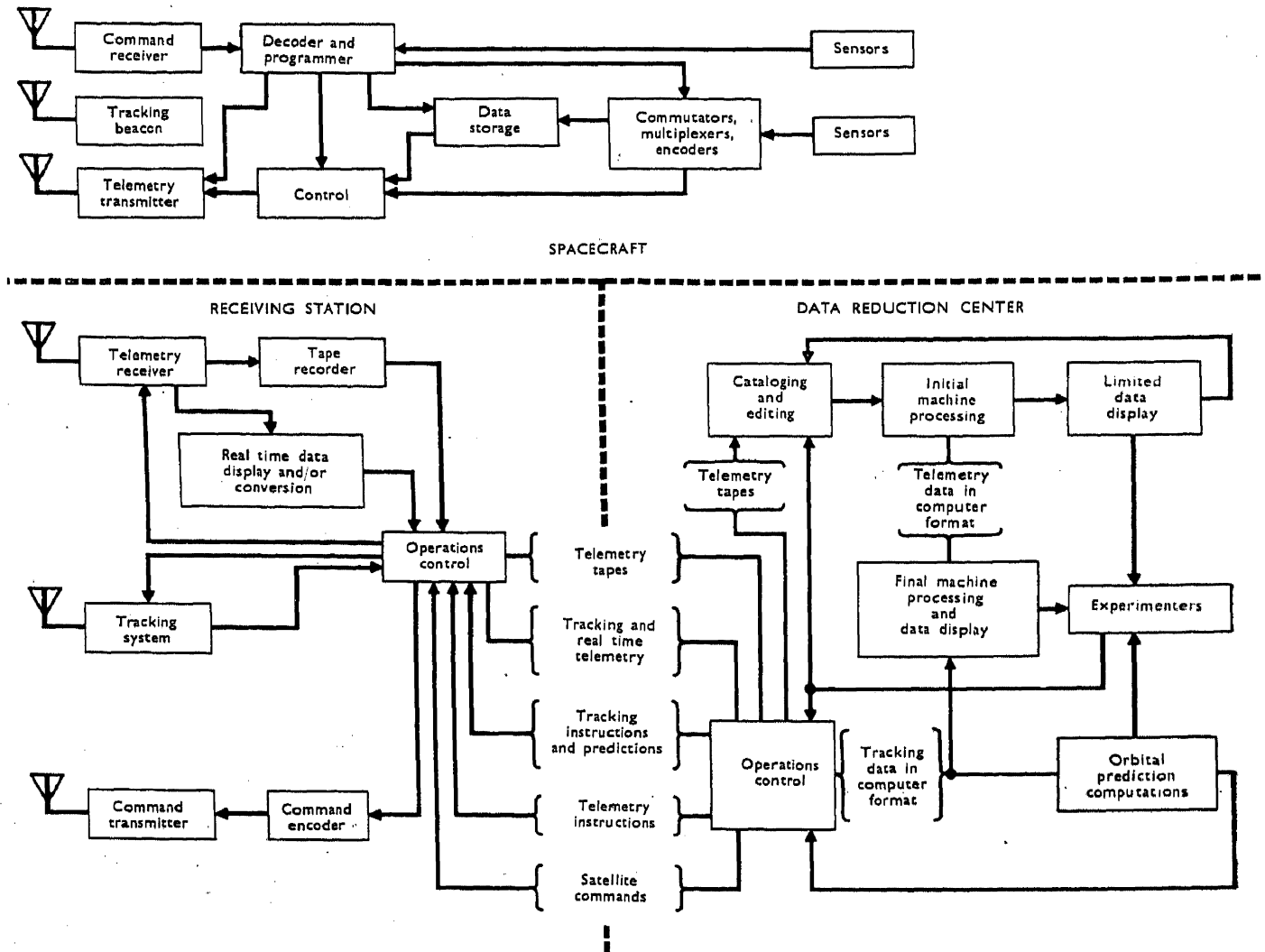


Fig. 27. - Information flow in spacecraft data, tracking and command systems.

star, for example), the power system, or the thermal control system, or *c)* control experiments by turning their power off or on, change scale factors, change absorbers, change viewing directions, etc. Some of the receiving stations possess transmitters and code generators to supply the necessary command signals. These commands may be transmitted according to a predetermined plan, or upon looking at a live or delayed data display.

3) Tracking. Stations which are established to receive spacecraft data usually have tracking systems as well. This tracking function may utilize parts of the telemetry receiver and antenna, as in the case of the Goldstone station, or it may be provided by a different antenna and receiver system, as in the case of the Minitrack interferometer tracking system.

4) Timing. Since it is necessary that specific experimental data be associated with position in space, it is necessary that accurate time be recorded with the received data. The use of a spacecraft clock will not obviate this need, since it will still be necessary to determine the spacecraft clock error by comparison with an accurate clock on the ground. The usual technique is to provide a stable, precise oscillator at the receiving station which is either continuously or periodically compared with the time from one of the standard time facilities, such as radio station WWV. This local time signal is recorded on the magnetic tape along with the received spacecraft signals.

5) Real-time display. For some spacecraft, it will be necessary to provide reduction of certain portions of the data in real time at some of the receiving stations (and perhaps ultimately at the operations control center by means of microwave or wire communications) in order that corrective action may be taken immediately in the event of a spacecraft subsystem malfunction or of a need to readjust an experiment. Little real-time display has been necessary so far, but, with the coming of the observatory-type spacecraft the need is rapidly developing for some real-time display at at least one central receiving station.

11. - Data reduction.

Once the telemetry receiver output is available on the magnetic tapes it is necessary to reduce these electrical signals into usable data. The magnitude of this task should not be underestimated. The number of data points received from a single satellite can be extremely large. From the single G.M. counter flown in Explorer I, which comprised about the simplest spacecraft data system conceivable, but where essentially no processing was done before transmission, of the order of 100 000 data points were received during the active satellite lifetime of 32 days (a data point is a number obtained from a detector which can be associated with satellite position, time, measuring interval, etc.). Other more complex systems have already yielded hundreds of thousands of data points, and the OGO spacecraft will be capable of providing $2 \cdot 10^{11}$ data points during its lifetime. Obviously the establishment of techniques for the handling of this volume of data cannot be assigned a low priority, to be taken care of after the launching is successful. It is also obvious that the time-proven techniques for reading data manually are not adequate for reducing spacecraft data, except for making data rapidly available for quick look investigations.

The usual procedure for the handling of the telemetry tapes is also shown in Fig. 27. They are forwarded to a central data reducing center located at

one of the NASA centers or, under certain circumstances, directly to an experimenter's laboratory. There they are partially reduced, and the results are recorded on a medium which can be directly used by the experimenters. These reduced data may take several forms, including computer tapes or cards, other magnetic tapes, oscillograms, or tabulations. The experimenters then further process these data so that they can conduct their analysis most conveniently.

The use of high speed computers for the initial reduction of the receiver output tapes and for the final reduction by the experimenters is becoming more and more necessary. Both of the two main data reduction steps, the separation of the signals from the noise, and the processing of these reconstructed data, can be accomplished by properly programmed general purpose computers or by properly designed special purpose computers.

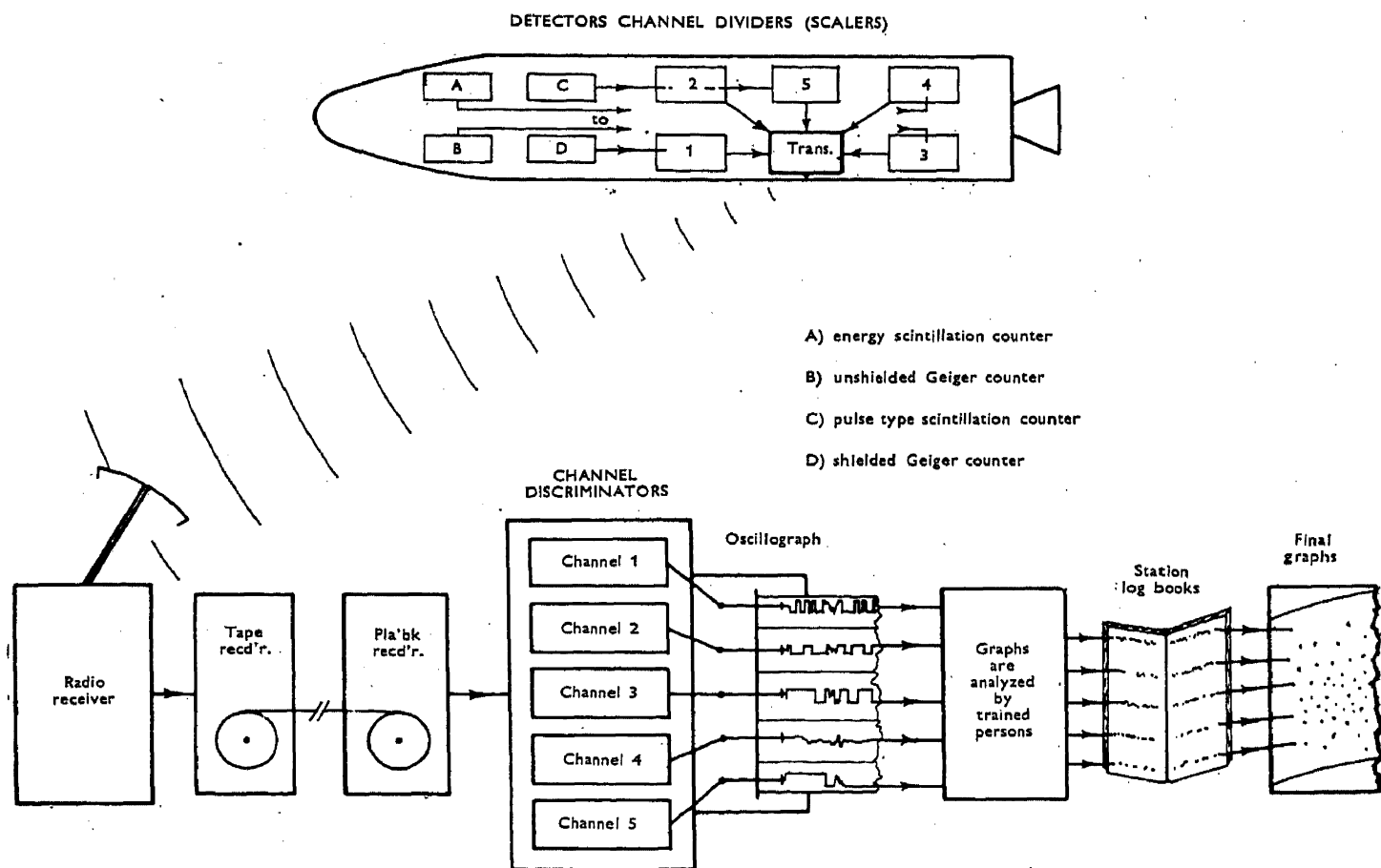


Fig. 28. - The Explorer IV manual data-reduction process.

A very simple data-reduction system is illustrated in Fig. 28. This is the one utilized in the early Explorer IV reduction program. The discriminators and moving pen oscillograph produced a paper chart, which was read by trained data readers. The manual operations required the expenditure of approximately six man years of effort to complete the reduction of about six

week's data. Subsequently, a machine reduction system was perfected which produced IBM punched cards in a single operation. It was discovered that the machine reduction method was considerably more accurate than the manual method.

An example of a representative present generation data-reduction system is the Explorer VIII system (Fig. 29). A similar system will be used for the S-3

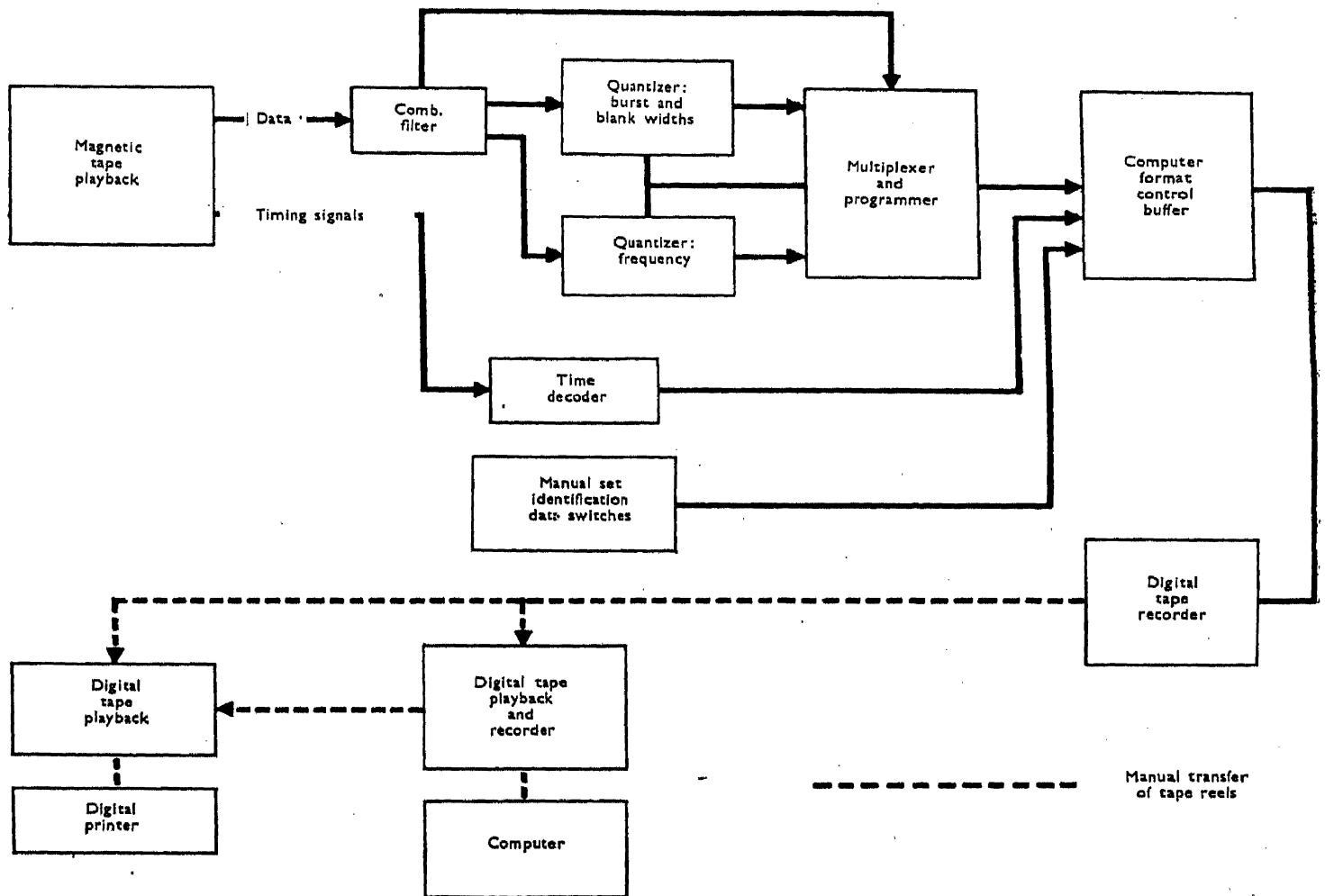


Fig. 29. - The Explorer VIII data-processing system.

and S-51 (U.K.) spacecraft. These spacecraft all make use of tone burst modulation, in which an audio frequency (5 kHz to 15 kHz) oscillator is gated on and off. In the S-3 and S-51 systems, the oscillator tone bursts of 10 ms are followed by 10 ms pauses. The frequency of the oscillator during the bursts is controlled by the data. A single oscillator can telemeter n digital bits, in which case there will be 2^n possible discrete frequencies. Alternatively, the burst can send analog data by making the frequency proportional to the signal amplitude. The data system includes a commutator, so that one burst is used for experiment 1, another for experiment 2, and so on. In S-3, the commutator has 16 inputs so that each frame is 16 bursts long. Frame synchroni-

zation is provided by making the first burst in each frame longer than the others. Several frames are necessary to send a complete number representing one detector reading.

In the data reduction system, the magnetic tape recorder output is fed through a 120-tooth comb filter, from which there are 120 output lines. The burst quantizer establishes the burst and frame synchronization. The frequency quantizer integrates the outputs of the 120 lines during the time that the bursts are present and selects the one having the greatest output amplitude. The outputs of the quantizers are fed to the decommutator, which reassembles the original data words. These words, plus the time signal, are recorded on a magnetic tape in computer format. These tapes may then be used for further data reduction, tabulation, or display by the computer, either at the central data reduction facility, or by the experimenter.

12. - Conclusion.

This paper has discussed a few of the problems involved in properly designing a spacecraft data handling system, including the initial conditioning of the signals from the sensors in the spacecraft, the transmission of these conditioned signals to the earth, and the processing of the data after receipt. Although it is not necessary that the experimenter understand all the details of many parts of the system, he should have a general understanding of the problems involved so that he can make the most efficient use of the spacecraft. He must be particularly familiar with many of the processes which occur at the two ends of the system, those which occur immediately following the spacecraft sensors, and those occurring just before the appearance of the data in final form at the output of the ground data reduction equipment, since these are places in the system where the basic form of the data is modified most strongly. It is necessary that the experimenter understand in what ways the data are modified by the complete data system.

Present spacecraft systems are capable of collecting very large quantities of data. The experimenter must remember that his role is to gain a greater understanding of the physical universe, not merely to act as an efficient gatherer of data. He must require that his instruments operate as efficiently as possible to assist him in this role. This implies that he will employ automatic data processing techniques to the limit of the currently available technology both in the spacecraft and on the earth for those processes which do not actually require his human judgement.

REFERENCES

- [1] J. E. WHITESITT: *Boolean Algebra and its Applications* (Reading, Mass., 1961).
- [2] G. H. LUDWIG: *Rev. Sci. Instr.*, **30**, 223 (1959).
- [3] V. E. SUOMI: *The Thermal Radiation Balance Experiment on Board Explorer VII*, Ch. 11 in *Juno II Summary Project Report*, Marshall Space Flight Center Report MTP-M-RP-60-1 (1960).
- [4] G. H. LUDWIG: *The Development of a Corpuscular Radiation Experiment for an Earth Satellite*, State University of Iowa Research Report SUI-60-12 (1960).
- [5] G. H. LUDWIG and W. A. WHELPLEY: *Journ. Geophysical Research*, **65**, 1119 (1960).
- [6] U. D. DESAI, R. L. VAN ALLEN and G. PORRECA: *NASA Technical Note* TN D-666 (1960).
- [7] C. E. MCILWAIN: *IRE Transactions on Nuclear Science*, NS 7, No. 2-3, 159, June-Sept. (1960).
- [8] J. W. FREEMAN: *A Satellite Borne Cadmium Sulfide Total Corpuscular Energy Detector*, State University of Iowa Research Report SUI 61-2 (1961).
- [9] H. L. RICHTER, Jr., W. PILKINGTON, J. P. EYRAND, W. S. SHIPLEY and L. W. RANDOLPH: *Electronics*, **32**, No. 6, 39 (1959).
- [10] G. H. LUDWIG: *The Instrumentation in Earth Satellite 1958 Gamma*, in IGY Satellite Report Number 13, World Data Center A, National Academy of Sciences (1961).
- [11] G. E. MUELLER: *Astronautics*, **5**, No. 5, 26 (1960).
- [12] *Reference Data for Radio Engineers*, 4th edition (Stratford).
- [13] E. RECHTIN: *Astronautics* **6**, No. 4, 37 (1961).
- [14] G. E. MULLER: *Proc. of the IRE*, **48**, 557 (1960).
- [15] A. G. SMITH: *Proc. of the IRE*, **48**, 593 (1960).
- [16] R. W. SANDERS: *Proc. of the IRE*, **48**, 575 (1960).
- [17] A. J. VITERBI: *IRE Trans*, SET-7, 3 (1961).
- [18] H. L. RICHTER, Jr., *Microlock: A Minimum Weight Instrumentation System for a Satellite*, Jet Propulsion Lab. External Publication No. 376, 15 Apr. 1957.

The British Space Research Programme.

E. B. DORLING

Department of Physics, University College - London

The first steps towards a British Space Research Programme were taken in the autumn of 1953 by the Gassiot Committee of the Royal Society. With the co-operation of the Royal Aircraft Establishment, a programme was evolved for the launching of measuring instruments into the upper atmosphere (up to 160 km altitude) by a specially designed rocket. Government financial support was obtained, the rocket known as Skylark was designed and made under the direction of the Royal Aircraft Establishment, and the Australian Department of Supply agreed to arrangements for launching at Woomera. Meanwhile in 1955, five University Groups started development work on experiments of their choice. The first flight trial of Skylark took place early in 1957, and since then many successful flights have carried scientific equipment to the 160 km altitude region. There is today a well established programme which is not only serving the needs of upper atmosphere research but is also including pilot equipment in preparation for earth satellite experiments. The Skylark rocket is powered by a Raven solid propellant motor giving a thrust of around 12 000 lb for half a minute or more.

Altitudes in excess of 160 km are reached with payloads of up to 65 kg. The whole rocket, which weighs about 1 ton, is $25\frac{1}{2}$ feet long and $17\frac{1}{2}$ inches in diameter. The 15° steel nose cone and the parallel instrument bay provide 5 cubic feet of space for the payload.

A later version of Skylark has an additional short burning time booster motor which gives a considerable increase in maximum altitude.

Telemetry equipment operates on a frequency of 465 MHz and can provide 24 channels of information. There are means available for parachute recovery of the instrument compartment, for ejecting part of the nose cone or for electrically insulating one part from another. It is also possible to carry some instrument sensors in the undisturbed region ahead of the rocket by mounting them on a 4 ft long spike projecting from the nose cone.

There is no doubt that Skylark is a reliable and versatile research rocket giving scope for a variety of valuable scientific measurements.

Before mentioning the scientific programme it is perhaps worthwhile stating some of the factors influencing our choice of experiments. In the first place it must be remembered that the advent of earth satellites, space probes and the like, while tremendously important and exciting, has in no way diminished the value of rocket research in the atmosphere below say 160 km. Rockets alone can examine a vertical cross-section of the atmosphere in a very short time. Satellites have a useful life only above the 240 km altitude mark; much of interest lies well below this altitude.

We are fortunate to have the use of the Woomera range in the Southern hemisphere; this enables our measurements to be made in geographic and geomagnetic latitude far removed from those of ranges in North America, North Africa and Russia.

Finally it is vital for the progress of geophysical and related sciences that systematic studies of the diurnal, seasonal and secular behaviour of the upper atmosphere and solar radiations should be made over long periods of time. The glamour of deep space probes and so on must not be allowed to overshadow the value of the patient systematic studies, for throughout the history of science, it is such studies that have yielded some of the greatest fruits.

To return to the Skylark programme:

During the past few years, the following research groups have carried out experiments:

University of Belfast: Measurement of upper atmosphere winds and temperature by the observation of sodium clouds released from rockets.

University of Birmingham: Measurement of ion spectra by the use of positive ion spectrometers ejected from the rocket on a long cable, and measurement of electron density by a dielectric method.

University College London: Data on upper atmosphere winds and temperatures have been obtained from observations of grenade bursts. Solar Lyman α -radiation and solar X-rays have been monitored by nitric oxide ionization chambers and emulsion detectors. Langmuir probe type instruments have been used to study sporadic E ionization and electron and ion concentrations in the ionosphere.

Imperial College London (Dept. of Meteorology): Observation by radar of the motion of resonant dipoles released from rockets have yielded data on upper atmosphere winds.

Imperial College London (Dept. of Physics): The variation of cosmic ray intensity with height has been measured.

One of the notable achievements of the Australian launching team at Woomera during 1959 was the successful launch of 3 Skylarks in 2 days for the international Rocket Week programme.

It is intended that the experimental programmes of the groups mentioned above will continue and expand, while other groups are gradually entering

the field; amongst them are the University College of Wales in conjunction with R.A.E., the Geophysics Department of Imperial College London, and the Royal Observatory Edinburgh.

Experiments soon to take place will be concerned with measurements of the geomagnetic field, observations in ultra violet wavelengths of the Southern sky, and ionospheric investigations using radio propagation methods.

In addition to these experiments being carried out in their own right in Skylark, we can expect this rocket to be used for giving test flights to equipment intended for the Scout satellite.

We now turn to the scientific programmes which have developed in the United Kingdom since the launch of Sputnik I in 1957.

It can be fairly claimed that the effort made in Britain to extract the maximum of scientific information from observation of the earth satellites launched by other countries has not only been on a considerable scale but has been very effective.

It is easy to allow the undoubtedly very great technological achievements of the satellites themselves to overshadow the valuable scientific work which has been and still is in progress, based on optical and radio observations. It would be unfortunate to ignore this work. Both the optical and radio observations vary in precision and origin. Some are from amateur observers using relatively inexpensive equipment but are none the less of considerable value, particularly in the prediction of satellite orbits of relatively low altitude. Others, from professional astronomers and missile tracking stations are of the highest precision, and some from the Jodrell Bank Radio Telescope are of course unique in that no other known equipment in the world can achieve them.

The methods used for optical observations vary; the simplest in operation is probably a plate camera in which a satellite track, photographed against the star background, is broken intermittently by a coded time mark. The use of kine-theodolites by trained operators has been very successful when weather conditions have been favourable and can give very precise results. At present stations so equipped exist in Scotland, West Wales, Southern England, and Malta. Other methods are under development, most of them with the aim of achieving the high accuracy of a kine-theodolite without the use of highly trained operators.

The methods for radio observations of satellites and space probes in this country are also diverse. We have already mentioned Jodrell Bank Radio Telescope, used for the reception of signals from deep space probes. Less sensitive, but adequate for earth satellites in near earth orbits, are the varieties of radio interferometers, such as that developed by the Royal Aircraft Establishment, and the American Minitrack system. A Minitrack equipment has been installed near the Radio Research Station, Slough. The output will be of use both for the American tracking network and for any purposes of our

own which may arise in the future. The data from precise optical tracking have already yielded new information about the ellipticity of the earth and about the effect of air drag on satellites. This latter study, leading to knowledge of the changes of air density at great heights and its variation with other effects such as solar radiation, is bound to play an important part in our developing knowledge of the upper atmosphere.

The study of radio propagation from satellites through the ionosphere is being energetically followed in this country and again is bound to lead to a greater understanding of this important region of the atmosphere. Finally tracking information is the raw data from which the satellite orbit is calculated and from which prediction of the time and direction of its appearance from any specified station can be made. While it might be expected that orbits well away from the earth can be calculated and predicted with precision for a long time ahead, this is certainly not so for satellites nearing re-entry, and passing through interesting parts of the upper atmosphere. Here the variability of air drag and the changing aspect of the satellite can cause large errors in prediction. The need for accurate up to date observations is here very evident. In the United Kingdom a satellite Prediction Service is operated by the Radio Research station at Slough. This station receives observations from many parts of the U.K., Europe, and the Commonwealth, and issues predictions to all who have a genuine need to know.

Many of you, I suspect, have a greater interest in the contents of satellites than in their external behaviour, so I will leave tracking at this point and turn to our plans and hopes for the direct use of satellites. As is well known to you, we have already arrangements with the American authorities whereby British designed and made instruments will be carried in American launched Scout satellites. There is also the possibility that British satellites might be launched by a rocket system based on the Blue Streak and Black Knight design. Whether or not this plan will take effect is at present undecided. I will mention the Scout plans first.

The offer by the United States to launch equipment designed and made in other countries was made to the members of COSPAR (the International Committee on Space Research) in March 1959. It was clear that Britain could gain much by taking up this offer. Our scientists had acquired 4 or 5 years experience with Skylark experiments, and some were in a good position to take the next step into satellite experiments. If it should be decided to launch British satellites with British rockets eventually early experience in Scout would be of great value. On the other hand, if no British launching took place, Scout seemed to offer the one certain way for our scientists to enter this new field. This is not to say that there are not many advantages in having a programme entirely under our control, but simply to recognize the fact that the very generous American offer could hardly fail to be of great value to us.

Accordingly plans were made for us to instrument several Scout satellites, to be launched within the next few years.

The Scout satellite launching system comprises a 4-stage solid propellant rocket booster. This is capable, in broad terms, of placing about 70 kg into an orbit of about 480 km altitude. There is of course a correspondence between payload weight and orbit. The Americans will undertake to launch the satellites carrying our payload, and in the first instance at any rate, will provide us with auxiliary services such as telemetry and power supplies. Launching will probably take place from a range on the East Coast of U.S.A. and we expect orbits to be achieved which will bring the satellite over the United Kingdom.

As usual, most Scout satellites will be « custom built » to meet the needs of particular groups of experiments, but broadly speaking the satellite can be thought of as a sphere some 2 feet in diameter, with antennae and other excrescences as needed. It will enter orbit spinning rapidly about its axis. This not only gives stability, but is advantageous in other ways, such as giving automatic scan for suitably mounted instruments. The availability of solar cells, giving power supplies adequate for a year's operation, and the opportunity to use some of the facilities of the American telemetry and data acquisition network, make Scout a very attractive vehicle for new-comers to the field. Nevertheless it is not difficult to envisage experiments for which it appears to be unsuitable.

The choice of British experiments for incorporation in Scout satellites naturally is restricted to those compatible with the facilities and with other experiments in the same satellites. It is also necessary that the scientific groups concerned should have prior experience with rocket-borne experiments. It is also of course important that an experiment should not mainly be a duplicate—unless there are special reasons—and should fit in sensibly with the international picture of space research. This can be ascertained through the channels provided by COSPAR.

For the first U.K. Scout payload a natural choice for the « core » of the experimental program was the direct study of the electrical properties of the ionosphere. Ever since the pioneer experiments of APPLETON, British scientists have played a leading part in studying this important region of the ionosphere.

Two research groups, one under Prof. J. SAYERS at Birmingham University and one under Dr. R. L. F. BOYD and Dr. A. P. WILLMORE at University College, London, have been developing the techniques for laboratory study of ionized gases such as the ionosphere.

Successful measurements have been made by both groups with instruments carried in Skylark rockets, and they are now preparing to extend the observations by using the Scout satellite. The University College group have

developed a probe for the measurements of electron temperature and electron density and two such probes will be mounted on a satellite, one on the base and one at the end of a boom. An identical boom on the other side of the satellite will carry the Birmingham University capacity probe, also, for the measurement of electron densities. A mass spectrometer for the determination of ion species will be provided by University College, who are also instrumenting the satellite with solar X-ray and Lyman- α intensity sensors; in this way the solar disturbances which affect the ionosphere can be monitored whilst the ionosphere itself is being studied. A solar aspect sensor will be carried to assist in the interpretation of the solar data.

In addition to these ionospheric experiments a cosmic ray experiment has been devised by Prof. H. ELLIOT of Imperial College London. The equipment will measure the flux of low energy cosmic ray particles of mass number greater than 6 and energy below 10^{10} eV. The sensor is mounted in the centre of the top dome of the satellite with the mass spectrometer above it.

This Scout International Ionosphere Satellite S-51, as it is known, carries four solar cell paddles which produce about 8 W, to charge chemical storage batteries which in turn power the equipment. The measurements are telemetered by a modified version of the Vanguard telemetry. An electronic encoder sub-multiplexes the various inputs and produces a series of F.M. bursts at a rate of 50 per second on a 136 MHz carrier, the frequency of each burst being an analogue of the voltage being telemetered. The transmissions are received at stations belonging to the Minitrack network and are recorded on magnetic tape. The tapes go to a data handling centre where the F.M. levels are converted to digital numbers and the data are prepared for computer analysis. It is expected that half a million data points will be collected each day.

A tape recorder is incorporated in the payload to enable data to be collected throughout a complete orbit at a much slower rate than with direct telemetry, and played out on command from a ground station.

The satellite is expected to have an apogee of about 1000 km, a perigee of about 320 km, and a lifetime of about 1 year.

Finally brief mention should be made of the plans now afoot for a European Space Research Organization. The United Kingdom is taking part in the work of the Preparatory Commission, and in the several working groups which have been set up by the Commission to discuss the proposed scientific programmes for both rocket and satellite launchings. It is very much hoped that before long this planning will give way to the actual setting up of a new small rockets launching range, a satellite payload engineering unit and a data and tracking centre and that Europe will embark on its own co-operative programme of space research.

Cosmic Ray Measurements in the U.S./U.K. Satellite S-51.

H. ELLIOT (*)

Imperial College - London

1. - Introduction.

The importance of the interplanetary magnetic field as a connecting link between many solar and geophysical phenomena has been repeatedly stressed during this course. We have also seen that although direct measurement of the field, using space probes, is now possible it will be a long time before a comprehensive picture of the large scale structure of the field is built up by this means. Meanwhile the cosmic rays can, in principle at any rate, provide a good deal of useful information about the average characteristics of the field. This is especially true now that it is possible to think in terms of protracted measurement of the primary intensity variations using satellites. By this means we can investigate not only the intensity variations of those primaries sufficiently energetic to produce progeny capable of penetrating to mountain altitudes or sea-level but also those lower energy primaries which are most sensitive to the magnetic fields we wish to measure.

In the present experiment it is proposed to investigate the cosmic ray spectrum by using a Čerenkov detector to measure the intensity of heavy nuclei ($Z > 6$) as a function of latitude. The energy spectrum can then be determined from the known values of the minimum energy which a particle must have in order to penetrate the earth's magnetic field at a given latitude. The advantages of measurements carried out in this way are: *a*) it avoids the invalidation of cosmic ray measurements by inadvertent detection of Van Allen particles; *b*) it avoids the introduction of uncertainties by albedo particles scattered back from the atmosphere.

The detector will sweep through the cosmic ray energy spectrum four times on each orbit, giving a virtually continuous check on its variation with time.

(*) The others members of the team concerned in this work were: Dr. J. J. QUENBY, Mr. A. C. DURNEY and Mr. D. W. MAYNE.

In addition, it is intended to carry out aeroplane surveys at the same time so that the intensity distribution in the atmosphere can be uniquely related to the primary spectrum. It is hoped by this means to obtain a sufficiently accurate relationship for the primary spectrum to be determined at any future time simply by an aeroplane survey.

A large and a small Geiger counter will be included in the instrument package. On those parts of the orbit where the satellite is outside the Van Allen belts the large Geiger will give a measurement of the energy spectrum of the primary cosmic ray protons and also bursts of solar protons in the same way as the Čerenkov detector measures the heavy particle spectrum. Thus it is hoped to be able to compare the proton and heavy particle spectrum as a function of time. Inside the Van Allen belts the large Geiger counter will saturate but the small Geiger counter will be able to count a much higher particle intensity. It is hoped to obtain data on the time variations of the trapped radiation from the small counter rate.

2. - Operation of the sensor.

The Čerenkov detector consists of a hollow perspex sphere, 4 in. in diameter, together with an E.M.I. 6097 photomultiplier of 2 in. in diameter which looks into a hole cut in the surface of the sphere. Čerenkov light flashes produced by cosmic rays in the wall of the sphere are detected by the photomultiplier. The pulse output of the tube is fed to a discriminator which only accepts pulses corresponding to primary nuclei of $Z > 6$ and rejects smaller background pulses due to lighter nuclei and the Van Allen particles. Output from the discriminator is fed to a chain of eight binaries and the contents of this store are sampled by both the data storage and direct telemetry encoders.

In flight calibration is provided since the Čerenkov counting rate is sensitive to the overall gain of the system. An α -particle source embedded in a scintillator will be used to provide standard light flashes for the photomultiplier. These light flashes are smaller than those produced by the passage of heavy nuclei through the perspex sphere and will not normally pass the discriminator. When the binary store is filled by cosmic ray counts, an electronic switch will reduce the attenuation of the photomultiplier output pulses to allow the α -particle pulses to trigger the discriminator. After these calibration pulses have filled the store, the switch will return the attenuation once again to the level for heavy nuclei. Thus a standard counting rate is provided for a fraction of the time on the telemetry so that the system gain may be continuously monitored.

The complete instrument package is located on the spin axis at the forward end of the satellite with the perspex sphere protruding outside the satellite

surface and the photomultiplier and associated electronics just inside. Primary heavy nuclei can arrive at the sphere from the forward 2π solid angle where the amount of shielding is in general less than 1 g/cm^2 but are absorbed by the satellite structure in the backward 2π solid angle. It is advantageous to locate the Čerenkov detector on the spin axis since the direction in which it sees cosmic rays only changes slowly in time with respect to the earth's surface and the geomagnetic field. The effect on the cosmic ray intensity of both the absorption by the earth and the deflection by the geomagnetic field must be taken into account in the analysis of the results.

The Geiger counters are placed in the instrument package with structure shielding of the order of 2 g/cm^2 from the outside in the forward direction. The effective length and diameters of the counters are respectively 5 cm and 2 cm for the larger and 1 cm and 0.3 cm for the smaller. An additional 1 mm of lead is put round the small counter. Both counters feed in parallel a chain of 13 binaries and the contents of this store are read periodically by both the data storage and direct telemetry encoders. If the counting rate of the large geiger tube exceeds capacity of the binary store, an integrating circuit turns it off with the result that only the smaller counter remains operative, so increasing the dynamic range of the system.

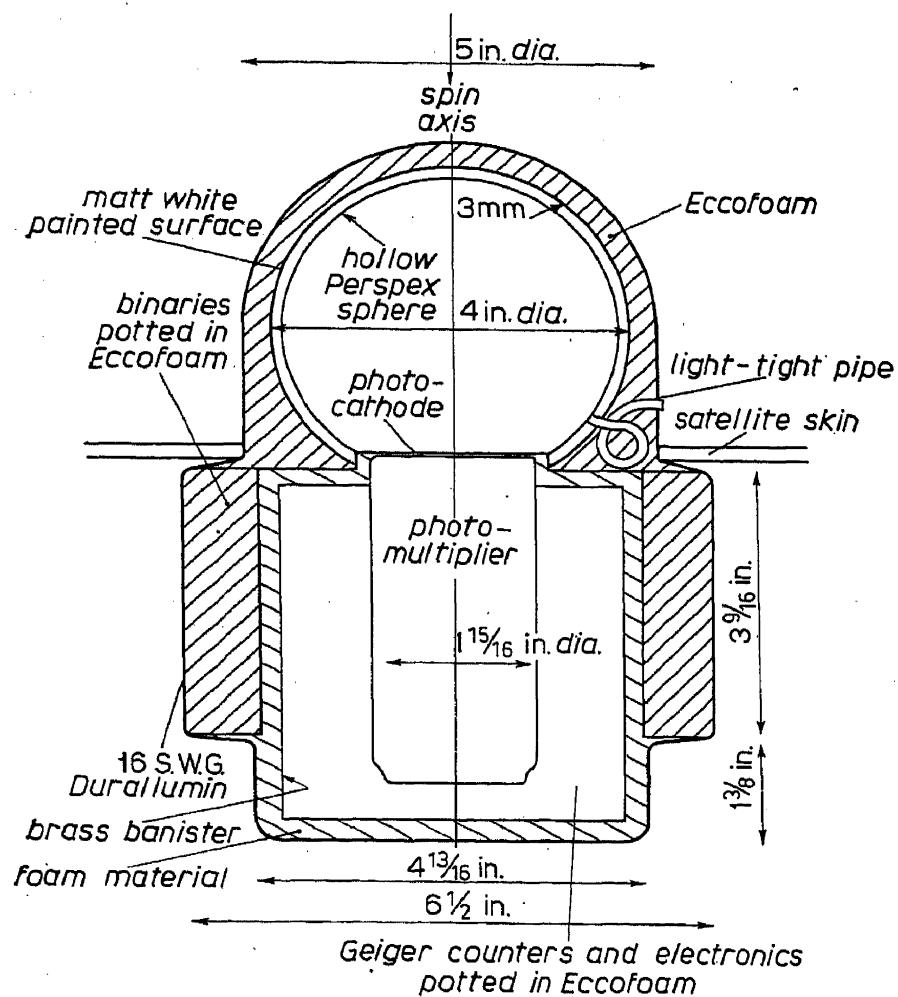


Fig. 1.

3. - Sensor structure.

A diagram of the sensor is shown in Fig. 1. The perspex sphere is contained in an aluminium dome, 5 in. in diameter and $4 \frac{5}{16}$ in. high, protruding outside the satellite skin, to which the stalk of the University College probe is attached at the top. Eccofoam serves as a packing between the sphere and the aluminium container. The sphere temperature must be kept below 70°C , the softening point of perspex. A light-tight air leak is provided between the interior of the sphere and the outside surroundings.

An aluminium cylinder $6\frac{1}{2}$ in. in diameter and 5 in. deep is attached under the dome inside the satellite and contains the photomultiplier, Geiger counters and electronics. The photomultiplier, $5\frac{1}{2}$ in. long, is placed along the spin axis and looks into a 2 in. diameter hole cut into the wall of the sphere. Some of the electronics, together with the two Geiger counters, are located in an annular space around the photomultiplier, and this whole is potted in eccofoam and placed in a brass shielding canister. The canister is floated in a foam material to reduce the vibration acting on the photomultiplier. The remainder of the electronics is situated in a further ring surrounding the floating canister, and this also is potted in eccofoam.

The estimated weight of the potted sensor, excluding the additional encoder circuitry, is 5.7 lb. The C. of G. is on the spin axis, $3\frac{1}{4}$ in. up from the lowest surface on the base of the electronics cylinder. The M. of I. about the spin axis is 0.2 lb. ft.^2 and about an axis through the C. and G. and perpendicular to the spin axis of the M. of I. is 0.09 lb. ft.^2 .

4. - Sensor electronics.

A block diagram of the electronic circuitry is shown in Fig. 2. Two input power supply lines are required, 200 mW at $(-9 \pm 1) \text{ V}$ and 100 mW at $(-6.5 \pm 1) \text{ V}$. The -9 V line is fed to a DC converter producing a very stable -6 V line which is used to supply the E.H.T. converter, the emitter follower, the discriminator, and the gate circuits.

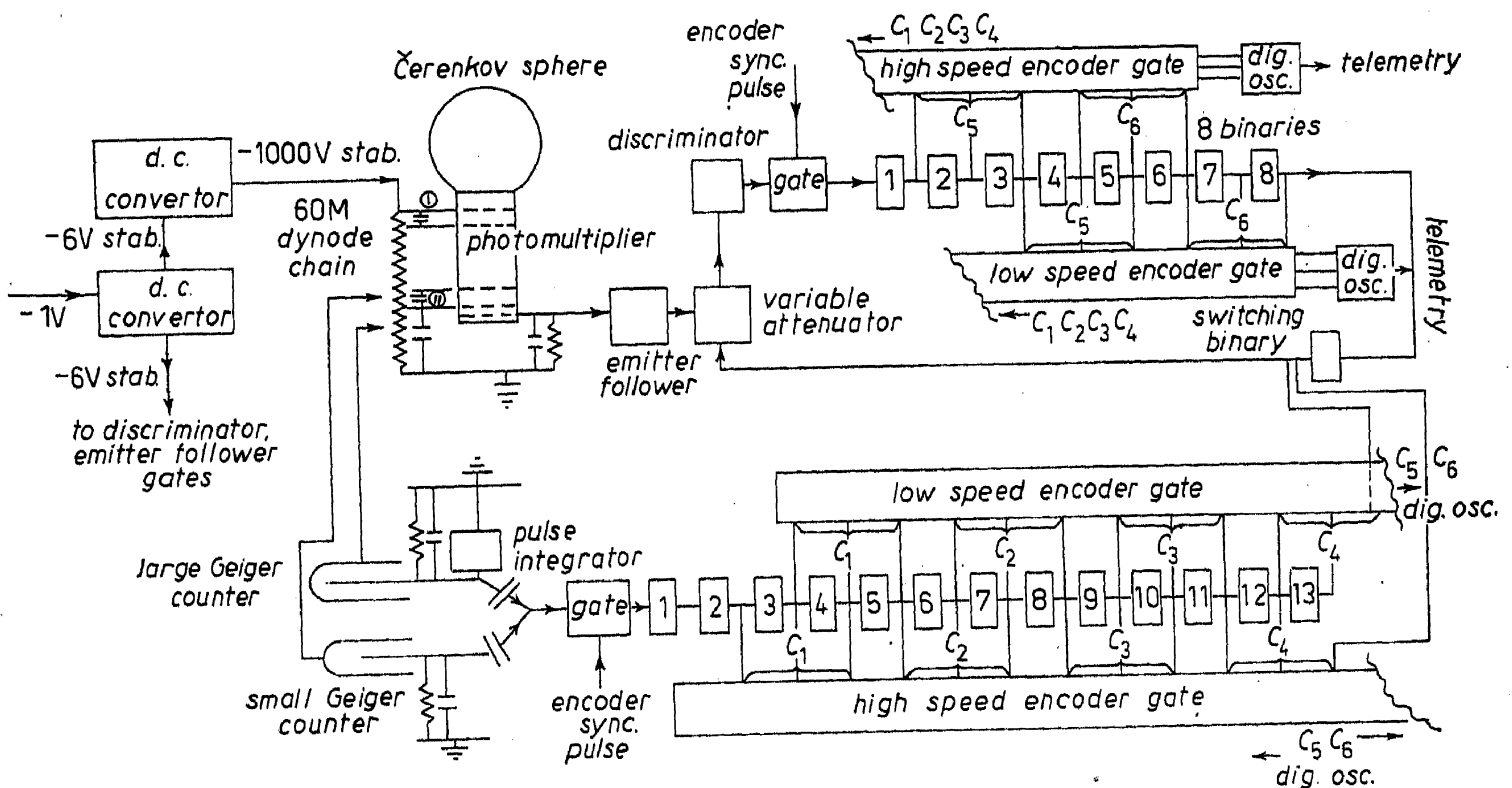


Fig. 2.

The DC converter incorporating a corona stabilizer supplies the $60\text{ M}\Omega$ dynode chain of the photomultiplier with -1000 V , and this E.H.T. voltage was designed to have a $\pm 4\text{ V}$ variation for a $\pm 1\text{ V}$ variation on the -9 V

TABLE I. - *Output voltage « V » under various conditions.*

Supply volts	Output with $100\text{ k}\Omega$ load (*)	Output with $50\text{ k}\Omega$ load (*)
5	3.5	3.1
6	4.3	3.9
7	5.0	4.7

(*) Load resistor D.C. connected between output and earth.
Note: the binary units will be operated from the unstabilized -6.5 V ($\pm 1\text{ V}$) supply.

input supply. Pulses from the photomultiplier are fed via an emitter follower and an attenuator to a discriminator. The discriminator level is designed to be stable to $\pm 3\%$ for $\pm 1\text{ V}$ variation on the -9 V input supply, and the attenuator is switched to permit measurement of either the calibration source or the cosmic ray rate.

The discriminator output goes via a gate circuit to the chain of eight binaries. Figure 3 shows the circuit of a single binary together with the size of the output waveform as a function of supply voltage. The outputs of the first six binaries are fed to the input gates of the high speed encoder, while the outputs of the last six binaries are fed to the input gates of the low speed encoder. When a particular information channel is to be sampled, the gate circuits feed three of the binary outputs to a digital oscillator which takes up one of eight different frequency levels, depending on the state of the three binaries. This frequency output is then put on the tape recorder in the case of the low speed encoder, and goes to modulate the telemetry transmitter in the case of the high speed encoder. Thus to read the binary store two channels are required on both the data storage and the direct telemetry systems.

While the low speed encoder is sampling the binary store, it is arranged that the gate circuit placed after the discriminator prevents any output reaching the first binary. This is done because the state of the binaries may change during the time taken to sample both binary sets. The gate circuit is driven by a pulse from the low speed encoder which has a quiescent level of $+6.7\text{ V}$ and goes to -2.7 V during the period when the state of the two sets of binaries is read by the digital oscillator. Zero current is drawn from the encoder at $+6.7\text{ V}$ and $10\text{ }\mu\text{A}$ at -2.7 V .

An electronic switch working from the output of the last binary will change the value of the two-position attenuator and successively feed the cosmic ray counting rate and the calibration pulse rate into the telemetry systems. When

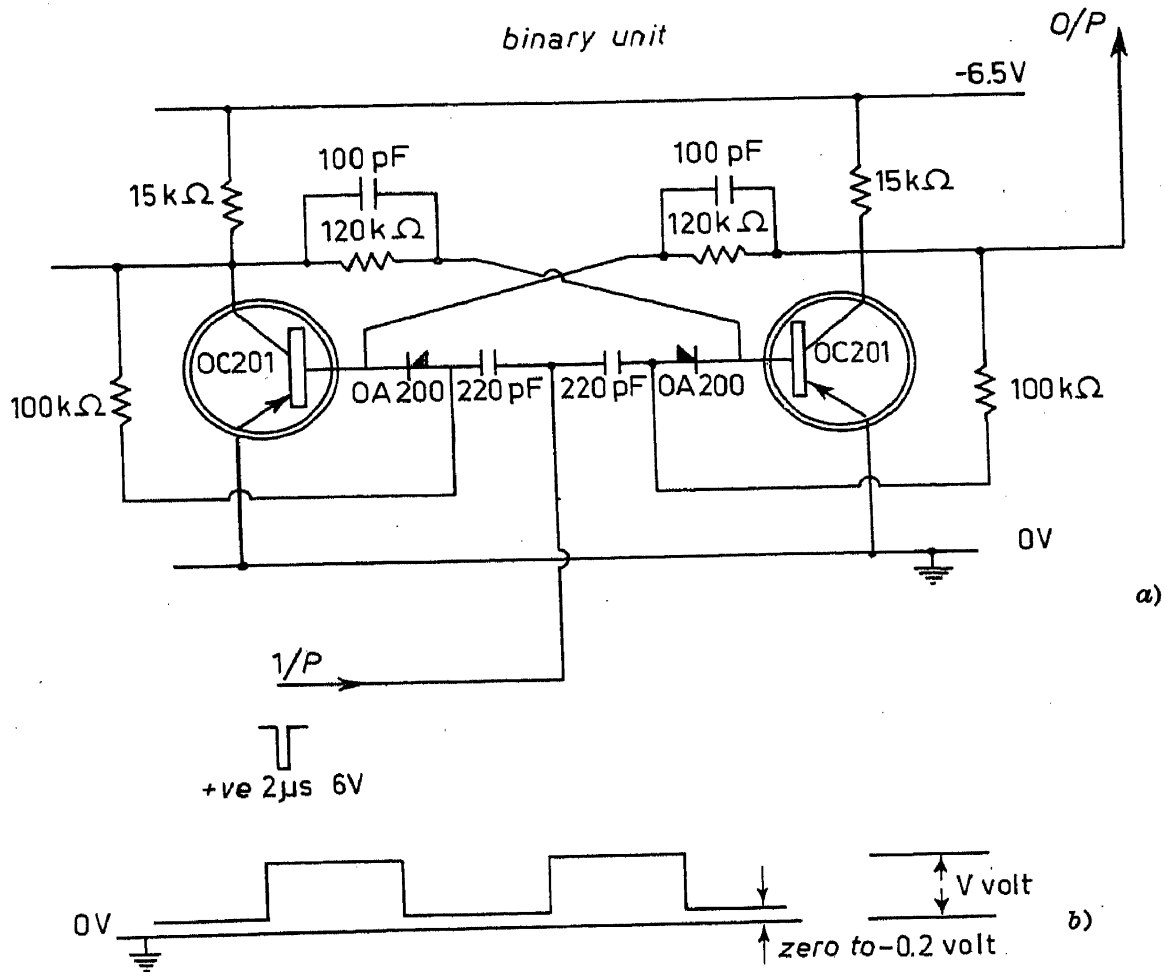


Fig. 3. - a) Circuit details. The output is a square wave. b) Output voltage « V ».

the binary store is filled, the flip-over of the last binary changes the state of a switching binary, the output of which determines the value of the attenuation. In order to tell which attenuation value is being used, the switching binary output is also fed to both the low and high speed encoders as one of the three binary elements in information channel C4, the other two elements being taken from the Geiger counter store.

Appropriate voltages for the Geiger counters are tapped off the dynode chain. The anode to cathode voltage of the larger counter is made to fall below the working voltage at counting rates in excess of that which the data encoder can handle. Pulse outputs from the two counters are mixed and fed via a gate circuit to a chain of thirteen binaries. Numbers 2 to 12 are read by the high speed encoder and numbers 3 to 13 are read by the low speed encoder. The design of the binaries and the method of sampling by the encoders are similar to the case of the Čerenkov channel. Thus four channels of both the

data storage and the direct telemetry systems are required. The gate circuit stops extra pulses being fed to the store while the low speed encoder samples the four sets of binaries and works from similar pulses to the Čerenkov channel gate.

5. - General considerations.

The basic requirement was to design the cosmic ray analysing equipment so as to form a compact package suitable for launching in a research satellite. The objectives in this work were 1) reliability, 2) economy in power consumption, 3) low weight, and 4) small size. The first two objectives were of paramount importance: the last two were subordinated to them. In addition, the equipment was required to work over a temperature range of -10°C to $+60^{\circ}\text{C}$.

The circuitry can be considered to comprise four basic sections, the power unit, the Čerenkov channel, the Geiger channel, and the data storage unit.

Silicon transistors, rectifiers and Zener diodes are employed throughout. Solid-electrolyte tantalum capacitors are used for high values of capacity and paper capacitors for medium values. For high working voltages and for small capacitances ceramic capacitors are used. The resistors are of the cracked, carbon high-stability type. A total of 121 semiconductors are employed of which 65 are transistors. The total power consumption is less than 300 mW.

The whole equipment is housed in an aluminium container measuring approximately 9.5 in. high and 6.5 in. in diameter. Its weight is 4.5 lb.

6. - Power supplies.

A number of factors influenced the selection of a power supply voltage. Since much of the circuitry is concerned with microsecond pulses, it is desirable to keep impedance low and a low supply voltage is helpful in this respect. Furthermore, a low supply voltage is conducive to power economy, a factor of a paramount importance in satellite instrumentation. Finally, the two factors, drift and circuit accuracy, have to be considered. Since these factors often depend on the stability of transistor parameters, it is important to keep the signal voltages high compared with the magnitude of the changes due to transistor variations. With these points in mind, the voltage selected from the power supplies available was 6 V. This, however, was an unstabilized supply, varying by about plus and minus one volt. It was immediately apparent that a far more stable supply was needed for the discriminator, emitter follower, and e.h.t. convertor and accordingly a 9 V line as well was selected to supply a voltage stabilizer giving a stabilized 6 V, controlled to plus or

minus 1%. The design of this stabilizer is largely conventional embodying a silicon Zener diode voltage reference, a long-tailed pair d.c. amplifier, and a 2-stage series circuit, *i.e.* four transistors in all. Compensation is also applied to produce a higher stabilization ratio than that which could normally be obtained.

The e.h.t. converter is fed from this stable 6 V line. It is a single-transistor convertor using a ferrite-core transformer and producing 150 V at the secondary. This secondary winding drives a voltage-multiplying Cockcroft-Walton ladder of eight silicon diodes to produce a 1200 V e.h.t. supply. The high oscillation frequency of 16 kHz was chosen for the convertor in order to keep the size of the capacitors in the ladder small. This high frequency also reduces the size of the filter required to isolate the e.h.t. convertor from the rest of the circuitry connected to the 6 V stabilized line. The filter consists of a 1.5 mH ferrite-core inductor and two 150 μ F capacitors included in the supply lead to the convertor. A low secondary voltage followed by voltage multiplication was chosen to enable conservatively rated silicon rectifiers to be used, and to ease the problem of secondary-winding insulation. The efficiency of the convertor, including the voltage multiplier, *i.e.* from 6 V input to 1200 V output, is 75%. A corona stabilizer is connected across the 1200 V line so giving a stabilized 1000 V supply maintained to within 0.4%. The photomultiplier dynode chain is connected across this 1000 V line; the Geiger tube supplies are derived from this chain.

7. - Čerenkov channel.

The photomultiplier generates negative going pulses of microsecond duration and of a few hundred millivolts amplitude at an impedance of about 500 k Ω . To reduce this impedance to a more convenient level, a two-stage emitter follower was designed, using high-speed silicon transistors, to give an input impedance in excess of 1 M Ω , with sufficiently low input capacity to handle these pulses with a minimum of distortion. The output impedance is a few kilohms.

The emitter follower feeds into a switched attenuator, which is simply a resistor switched in and out of circuit by a transistor. Since in one condition this transistor is required to present an open circuit to fast pulses a very low collector capacitance is called for and the type of transistor used was chosen with this in mind.

The discriminator is of the emitter-coupled monostable type and a single resistor is employed to set the discrimination level. This is necessary since each equipment is individually calibrated and adjusted to allow for such variations as the photomultiplier gain.

The pulse amplifier and gate consist of a grounded-emitter limiting amplifier, with its input shunted by a transistor which is switched, via another transistor, by the gating pulse.

8. - Geiger channel.

Two Geiger tubes are used, one of high and the other of low sensitivity. Their outputs are connected in parallel and the circuit is arranged so that in high radiation regions the high sensitivity tube passes sufficient current into an integrating network to reduce its supply voltage to below the operating plateau. The two inputs are fed directly into a pulse amplifier and gate similar in design to those in the Čerenkov channel.

9. - Data storage unit.

The data storage section is composed of 22 binary stages, eight for the Čerenkov channel and 14 for the Geiger channel. These are of conventional form but designed to operate with maximum power economy from one supply line only, *i.e.* the — 6 V stabilized supply, each binary stage consuming 2.4 mW. Here, to some extent a compromise is made between economy and maximum operating speed. Reduction of current consumption leads to higher impedances and with given transistors a lower maximum speed results. Most of these binaries will operate up to 30 kHz, although 10 kHz is considered to be the reliable maximum for production modules. Since the maximum pulse rate expected in this equipment is 1 000 per second, there is an ample safety margin.

10. - Mechanical considerations.

It is considered that, when using conventional components, the greatest component density is achieved by using a « sandwich » form of construction. In this arrangement, the components are strung between two parallel printed-circuit boards, care being taken to choose components of approximately equal lengths. This form of construction has been adopted as far as possible in this equipment.

Since there are 22 identical binary units, these were treated as individual modules in « sandwich » construction, each measuring approximately 1 in.sq. $\times$ $\times \frac{3}{4}$ in. deep.

The basic mechanical arrangement of the analyser is determined by the photomultiplier which must be resiliently mounted. Further more, since it was considered undesirable for high voltage leads to pass through the resilient material, the d.c. convertor and associated circuitry have been mounted with

the photomultiplier. Long leads at high impedance are also undesirable and the emitter follower is likewise included with the photomultiplier as are the Geiger tubes and their immediate circuitry. All these units are housed in a perforated brass cylinder divided into compartments to screen the high impedance circuits from the e.h.t. convertor. The whole is floated in neoprene foam. The remainder of the circuitry is embodied in three separate toroids around the cylinder, one of « sandwich » construction and two carrying the binary modules.

A detachable dome assembly, housing the perspex sphere, fits over the cathode end of the photomultiplier.

All the electronic circuitry is encapsulated in « Eccofoam » a proprietary rigid foam-in-place polyurethane resin which contributes to the rigidity of the structure.

Aluminium was chosen as the material for the container and since the instrument took a cylindrical form, of three different diameters, with a hemispherical end, spinning was selected as the method of manufacture. This avoids welded or rivetted seams and produced a continuous metal structure without the expense of machining from the solid.

The external finish above the satellite skin is a high polish ready for the application of the correct thermal coating.

Note added in proof.

In the final flight version of the cosmic ray sensor, the large G.M. tube was omitted together with the calibration source. The calibration mode was used to provide a second discriminator level for the detection of particles with $Z \geq 2$. Thus the sensor first counts 256 particles with $Z \geq 6$ and then switches over and counts 256 particles with $Z \geq 2$, and so on.

Research and Research Plans at Kiruna Geophysical Observatory.

E.-Å. BRUNBERG

The Royal Institute of Technology - Stockholm

1. - Introduction.

The Kiruna Geophysical Observatory was formally opened in July 1957 after a decade of planning. It is situated in the northern part of Sweden in the town of Kiruna at the geographical co-ordinates, lat, 67.8° N, long. 20.4° E and thus in the auroral zone (Fig. 1). The observatory belongs to the Royal Swedish Academy of Science.

Dr. B. HULTQVIST has been the director since the start and about 20 people are attached to the observatory at present. It is already fairly well equipped and enables geophysicists interested in arctic phenomena to carry out research under the most favourable conditions.

The observatory consists of a main building and a series of smaller buildings for housing various types of instruments. The observatory covers an area of 22 km^2 , which provides good possibilities for future expansion, including the establishment of observing sites, where measurements over base lines of several kilometers are desirable. The geographical location is particularly significant with respect to the Geophysical Institute in Fairbanks, Alaska, because the two observatories are approximately 180° apart in longitude. This circumstance makes it possible for the two laboratories to undertake certain important types of trans-polar research projects on a co-operation basis.

The main building of the Kiruna Observatory contains a series of laboratories, dark-rooms, administrative offices, library and seminar room, dining room, living quarters for investigators, a well-equipped mechanical workshop and other necessary facilities. General laboratory services are available in each of the rooms.

On the top floor of the main building is a glassed-in laboratory which is

also connected to the central heating. Above this laboratory is an open deck on which equipment can be mounted requiring all round viewing conditions. The mounting facilities are provided with supports which go down to foundations and are not connected with the building itself.

After this general presentation of the observatory I will tell you a little about the research work, which is going on there. Much of this work is collected under the following heading: «Electro-magnetic phenomena in the upper atmosphere or at the earth's surface which are correlated with each other and strongly depend on the conditions on the sun.»

The following table gives a survey:

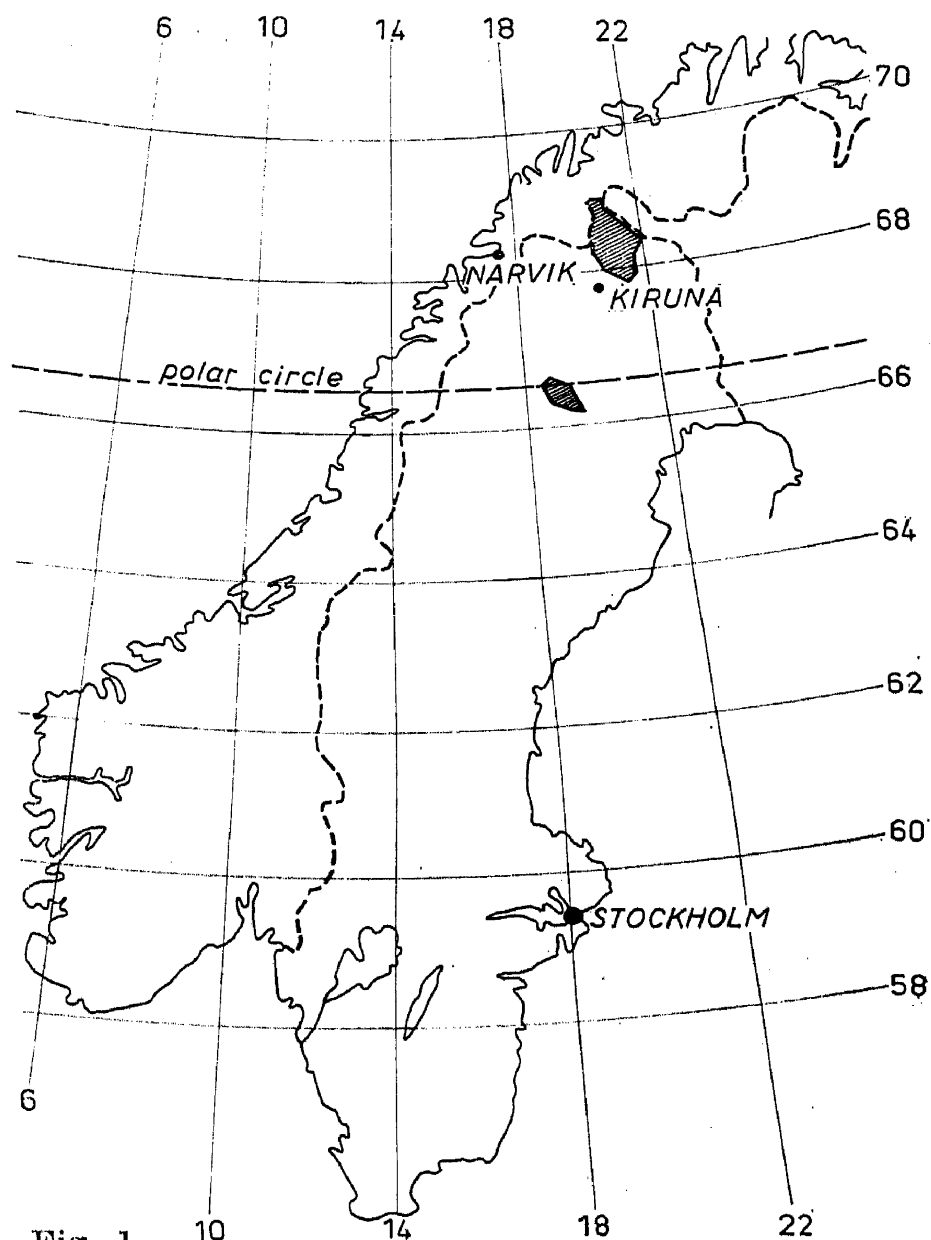


Fig. 1.

Geomagnetism	Auroral investigations		Ionosphere	Radio wave propagation	Cosmic rays
	Optical	Radio			
1. Registration of the 3 components	1. All sky camera	Reflexions	1. Ionosond	1. Alaska-Kiruna, short-wave	Meson telescopes
2. Magnetic variations in the audio frequency range	2. Spectrophotometer		2. Riometer	2. Rugby-Kiruna	
			3. Radio-astronomic swept-frequency interferometer	3. Transmission from satellites	

All the above mentioned investigations are carried out by automatic recording on film, paper or tape.

2. - Geomagnetism.

Measurements of the earth's magnetic field in the three components have been made since 1950 at a place not far from the new observatory and is now continued in the new buildings. The result is automatically coded on paper-tape, which every day is transmitted to Stockholm on a teleprinter and there evaluated in an electronic computer.

3. - Auroral observations.

Aurora is observed in Kiruna by optical methods as well as by reflexions of radio waves. Among the optical methods is taking photos with an all-sky camera and spectrophotometric measurements. The all-sky camera records every night during the darker part of the year. It was constructed by Stoffregen and started to operate during the geophysical year 1957-1958.

The apparatus consists basically of a spherical mirror, which projects the whole hemisphere via a plane mirror into a 16 mm cine camera. One picture is taken every minute with an exposure time of about 5 s. The films are not analysed in Kiruna, but sent to the different world centers for aurora, that were set up during the geophysical years, one of which is in Stockholm.

The spectro-photometer consists of a system of plane mirrors which is rotated and scans different parts of the sky. The light is collected by a parabolic mirror and an optical system, passes through a filter which selects the wanted frequency range and then enters a photo-multiplier. The signals from the photo-multiplier are finally recorded on tape or on an oscilloscope. The photometer is used to investigate auroral spectra with the chief stress laid upon rapid time variations. Absolute measurements of auroral intensities, studies of rapid variations of the signal (up to 20 kHz) and determination of the luminosity curves in six different spectral ranges are done.

Aurora is also studied by reflected radiowaves from the discharge. About 600 km south of Kiruna there is an F-M transmitter working at 90 MHz. If the antenna at the observatory is directed towards the transmitter, no signals are received because the transmitter is below the horizon, but when during aurora the antenna is directed towards north, signals are received. The reflected wave is modulated by the auroral discharge and a study of the «fading» can tell us something about the physical properties of the auroral mechanism.

4. – Ionosphere.

Investigations of the ionosphere with a Radiosonde in the traditional manner, started in Kiruna as early as 1948 and are now carried out at the new Observatory.

Rio-meter (relative ionospheric opacity) is a receiver for cosmic radio noise and measures the absorption in the ionosphere of this noise at about 27 MHz. The apparatus has proved extremely important when studying ionisation phenomena in connection of solar outbursts.

A *radio-interferometer* for studying radio-astronomical sources consists in principle of two receiving antennas at a certain distance from each other.

For a given frequency, the maximum intensity is recorded in principle when:

$$n \cdot \lambda = a \cdot \cos \alpha,$$

where n is a whole number and λ is the wavelength see Fig. 2. The antennas are helices.

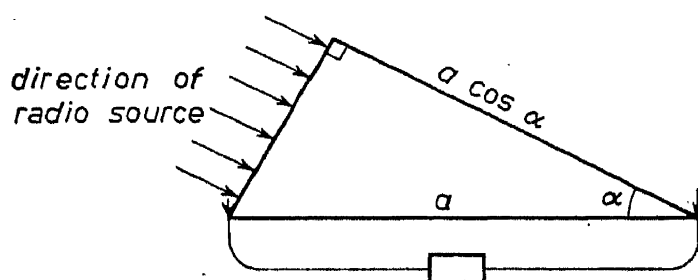


Fig. 2.

A swept frequency interferometer means that the frequency is continuously varied (in Kiruna between $(35 \div 65)$ MHz). Such an interferometer has several advantages in comparison with a fixed frequency instrument as for instance when determining the frequency dependence of the measured emissions. The swept frequency also permits accurate measurements of rapid movements of the source.

Radio propagation investigations over the polar cap are performed in collaboration with the Geophysical Institute, College, Alaska. Thus in Kiruna transmissions on 12, 18 and 24 MHz from Alaska are continuously recorded. The signal strength of the Rugby transmitter in England (16 kHz) is also recorded.

Cosmic ray intensity variations, are recorded by two directional Geiger-Müller tube telescopes pointing north and south as well as by the international standard cubical telescope. I may, in this connection mention, that at present there are two cosmic ray intensity recording stations in Sweden, in Kiruna and in Uppsala near Stockholm. Dr. SANDSTRÖM in Uppsala is responsible for these stations. The important task is to compare the time variations, measured at the stations after corrections for the geomagnetic deflection of the primary particles and the atmospheric influence, thus trying to find the strength and direction of any cosmic ray anisotropy that exists. There are

also measurements from Spetsbergen with directional telescopes available for the period August 1957-April 1959 and it will be most valuable to see if these three stations give a consistent picture of the cosmic ray conditions around the earth during this period.

Now we have gone through the list. It is naturally of great advantage to have at the same observatory, such a wide field of different registrations of geophysical phenomena. However, there are quite a number of other measurements carried out as for example:

Meteorological readings of air temperature, air pressure, humidity, rainfall, time of sunshine and temperature at different depths in the earth's surface.

Seismographic registrations are favourably carried out in Kiruna, where the rock is especially good for this purpose. On the seismographic recordings it has been possible to distinguish atomic bomb explosions as well as the « noise » from the heavy sea along the Norwegian coast.

Air sampling for measurement of the artificial radio activity in the atmosphere as well as certain chemical investigations.

This was the survey I wanted to give about this research work going on at Kiruna Geophysical Observatory. The plans are to continue this program for a long period of time, which is of great importance to do, because the registrations of geophysical phenomena with the same instruments have to be carried out during long periods, (of the order of a solar cycle) to give the most complete information.

One activity, which probably will be of great importance for the observatory in a near future is space research by means of vertical sounding rockets. The present Swedish plans are to begin launchings from a rather small range near the polar circle this autumn. The first experiments will be meteorological investigations at heights between (60 ÷ 90) km. Artificial clouds will be created at these heights and their movement observed by optical methods. Thus giving information about wind velocities and turbulence. It is a rather simple experiment suitable for a start in rocket research but we hope it will give valuable information especially about the physics of the noctilucent clouds which are frequently observed at these latitudes during some weeks during the spring and the autumn. Small american rockets will be used.

The present Swedish range is, however, too small for future experiments at higher altitudes and negotiations are going on in order to make it possible to have a launching range in the vicinity of Kiruna (Fig. 1). Such a range may also be of importance for the European Space Research Organization (ESRO), which is now being planned.

There is a whole range of interesting investigations to carry out by means of rockets in the auroral zone. In particular, the measurement of the flux, energy- and directional distribution of electrons and protons as a function of

altitude. For example, during polar cap absorption the proton flux should be measured from 30 km and upwards. Apparatus for this is now under development.

Among other plans are investigations of excitation and ionization mechanism in the auroral layers by injection of a suitable testgas in the relevant altitude range and spectroscopic investigations of its emission from the ground.

Before I end the presentation of Kiruna Geophysical Observatory and the research work there, let me point out that Dr. HULTQVIST and his co-workers are very glad to see physicists from other countries who are interested in doing research work at the observatory. Kiruna has probably one of the best situated geophysical observatories of the world north of the polar circle. The climate is good, there are railway and good road connections and daily services by air from Stockholm, which is reached in four hours.

Programs of Radioastronomy in Italy.

M. CECCARELLI

Istituto di Fisica dell'Università - Bologna

1. - Solar research.

The largest part of solar radioastronomy research is carried out at present in Italy at the Astrophysical observatory in Arcetri. The observatory is equipped at present with three instruments; an array of four helical antennas working at 220 MHz, a five metre parabolic mirror at 1400 MHz and a two metre mirror at 10 kHz. All three instruments are connected to « total power receivers » and are engaged in the continuous survey of the Sun. In addition to these there has been recently mounted a two element interferometer, also working at 220 MHz, for a better angular discrimination of active regions of the Sun.

Solar research is also performed at the Bologna radio observatory by the use of the existing radio telescopes which consist of a parabolic cylinder mirror (7×110) m large, and of a ten meter diameter parabolic mirror, both working at 327 MHz. These instruments are used daily to obtain transit records of the Sun and to study occasional phenomena such as the occultation of the Crab Nebula radio source by the solar corona.

For the future, the Florence group envisages the possibility of completing the equipment for solar studies by the construction of a radio spectrograph. The antenna for this and other instruments will probably be a ten metre parabolic mirror having mechanical tolerances of the order of one millimetre and which should be delivered by an Italian factory within about one year.

2. - Galactic and extragalactic research.

The major activity of the Bologna radio astronomy group is at present to design and construct a large antenna together with the corresponding receiving system, particularly suited for high resolution studies of galactic and extra galactic sources.

The form and dimensions of the instrument, operation frequency, receiving system etc., have been fixed by various considerations involving both the economic side of the programme, the international standpoint of this sort of research and the necessity of equipping the country with a high class instrument within a reasonably short time.

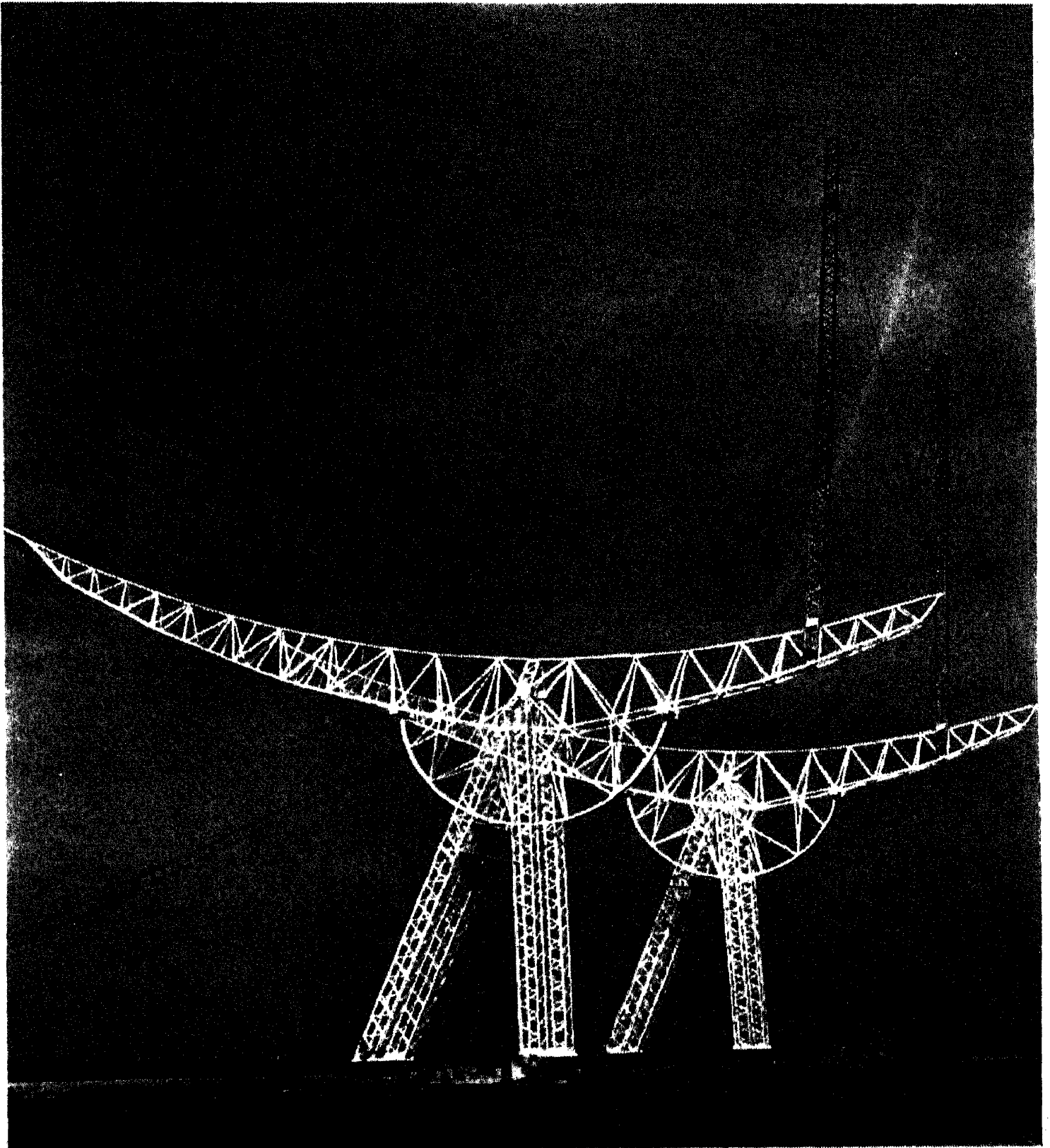


Fig. 1.

The chosen design has been that of a cross antenna, each arm having a length of about 1200 m and an area of about 30 000 m², operating at a frequency of 408 MHz. The arms will be made by parabolic cylinder mirrors; a type of construction which allows such large dimensions at a comparatively low price. Fig. 1 is a photograph of the first two parabolic sections of the E-W arm. Fig. 2 is a model of the N-S arm.

The E-W arm is a single cylinder of steel wires supported by 50 parabolic sections (Fig. 1).

The N-S arm (Fig. 2) will be formed by 128 parabolic cylinders which can be rotated in the meridian plane and are spaced so as to avoid mutual shadowing when they are pointed away from the zenith. The beam of such a discon-

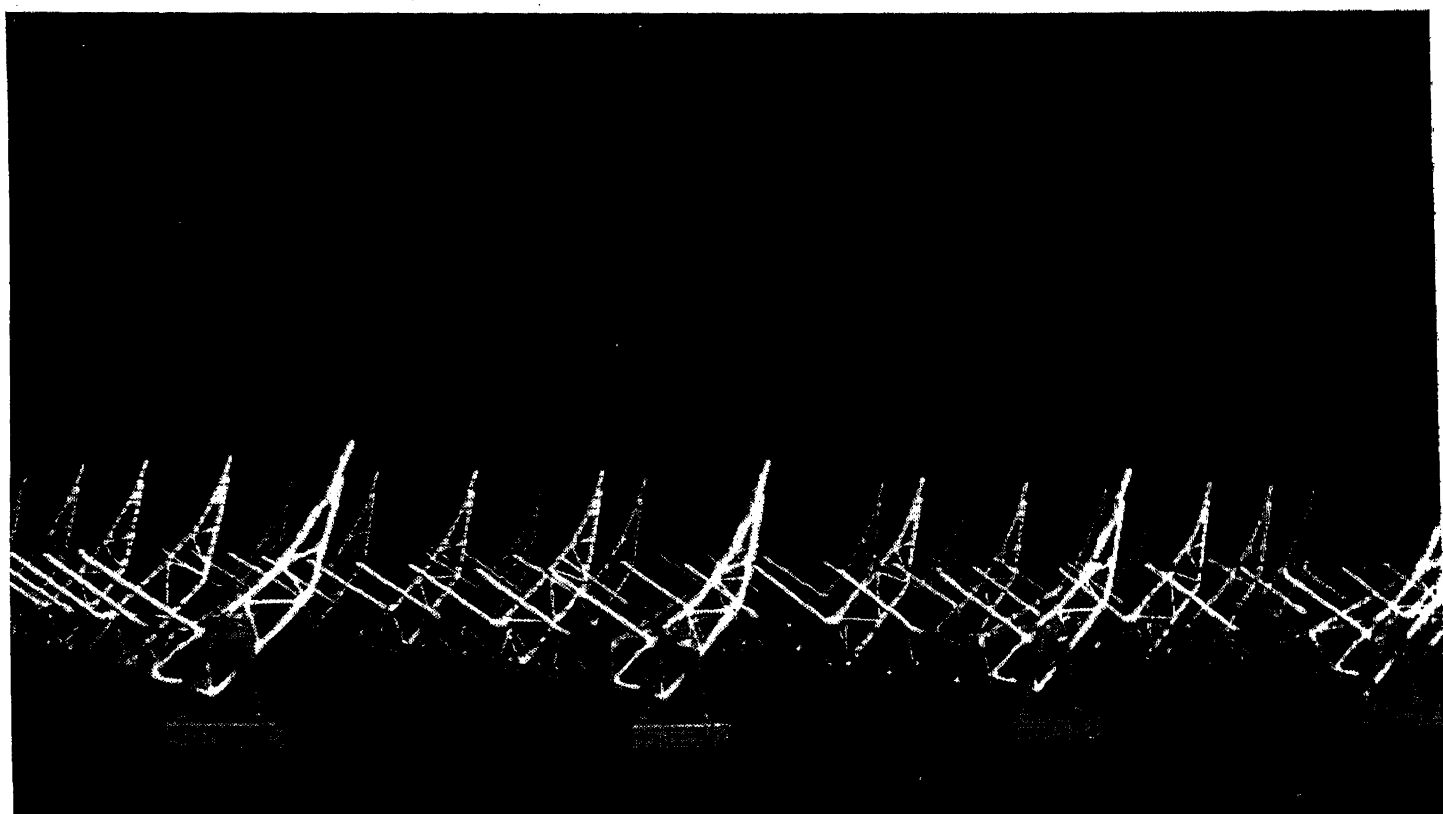


Fig. 2.

tinuous structure will of course have secondary maxima; these are however eliminated by multiplying the N-S pattern by that of the E-W arm, which is narrower than the fringe spacing. A discussion of the electronics is entirely outside the purpose of this lecture; we might simply say that the design foresees the use of 32 parametric diode up-converter preamplifiers, distributed along the two arms of the cross, and the simultaneous recording and electronic preprocessing (parasitic cut-off, integration, etc.) of at least 15 independent beams each covering an angle of about 4×4 minutes of arc. Detailed research programmes with such an instrument, which will be operating only in about two years, will

of course be largely affected by the evolution of research in this time. The fundamental objectives remain: an accurate, error-free mapping and flux measurement of very faint radio emitters and structural studies of nearer objects. The first topic will be a source of data on statistical properties of radio galaxies, cosmological theories and the formation of radio galaxies; the second should give a deeper insight into the processes of radio emission, particle acceleration etc., by extra galactic and galactic objects.

International Agreement as a Vital Factor in Radio Astronomy and Space Research.

H. C. VAN DE HULST

Leiden Observatory.

1. – Introduction.

International organization exists in sport, in business, and in most professions, and since it is generally unavoidable in the present world, it must also exist in science. The interesting problem is not why it exists but whether active scientists (like us) should become involved. In sport, international matches are played but the players themselves do not interrupt their training for travelling around to discuss rules, to fix dates, and to settle disagreements. This job is done by people who are not, or are no longer active athletes. If this example were a good precedent for science, the conclusion would be: let active scientists occasionally go to a symposium, but let them *strictly avoid* losing time in organizational meetings. I suppose this attitude has a natural appeal to you, and needs no further elaboration.

Now I am going to defend the opposite point of view, namely: *all active scientists* in our fields (radio astronomy, space research, etc.) should be aware of the need for international agreement, and *some of them* should spend considerable time in organizational work to reach such agreement.

Some organizations, to which I shall refer, are listed by their common abbreviations in Table I. It is easier to list their names than their functions, because (as in any organism) these functions change continuously. There are also certain ailments (including arteriosclerosis) and very complicated mutual relations via official channels and by «induction».

I can get to the point most quickly by starting at the difficult end, namely by listing some functions which these organizations have to perform. It will become clear, I hope, that some of these functions must be exercised at a level of agreement between nations (= governments), whereas others should be agreed among colleague scientists.

For instance it may be clear at once (after looking up the full names in Table I) that WMO and IAMAP, though they deal both with scientific *and* organizational problems in the same field Meteorology, may quite well exist side by side with different functions and different emphasis on these problems.

TABLE I. — *Some organizations referred to in this talk.*

<i>World-wide governmental</i> (national representation by governments)	
UNO	United Nations Organization
UNESCO	United Nations Educational, Cultural and Scientific Organization
WMO	World Meteorological Organization
ITU	International Telecommunications Union
CCIR	Consultative Committee for International Radio
IFRB	International Frequency Registration Board
ICAO	International Civil Aviation Organization
<i>World-wide nongovernmental</i> (national representation by Academies and similar bodies)	
ICSU	International Council of Scientific Unions
IAU	International Astronomical Union
IUPAP	International Union for Pure and Applied Physics
URSI	International Scientific Radio Union
IUGG	International Union of Geodesy and Geophysics
IAMAP	International Association for Meteorology and Atmospheric Physics
IAGA	International Association for Geomagnetism and Aeronomy
CSAGI	Special Committee for the International Geophysical Year (discontinued)
COSPAR	Committee on Space Research
IUCAF	Inter-Union Committee on the Allocation of Frequencies
CETEX	Commission on the Contamination by Extra-terrestrial Exploration
<i>Regional</i> (Western Europe)	
CERN	Centre Européen des Recherches Nucléaires
COPERS=ESRO	(Preparatory commission for the) European Space Research Organization
ESO	European Southern Observatory

Groups of colleagues, with contacts through correspondence and visits have always existed, and have gradually grown into the « International Scientific Unions », such as I.A.U., U.R.S.I., I.U.P.A.P., etc. A typical assembly of these Unions comprised about 50 people in 1920, but is more nearly 1000 people in 1960. They have developed an elaborate substructure of commissions. In I.U.G.G. there are the independent « associations », two of which are mentioned in the Table. They also have a superstructure in I.C.S.U., which confederates 13 International Scientific Unions and about 50 National Academies.

2. - Personal contacts.

This was the chief function of the Unions, because co-operation in science presupposes mutual understanding, and this in turn presupposes personal contacts. Many people feel that the present assemblies of the Unions are getting too large, and experiments are being made with other groups, such as this Summer School, and symposia of many kinds. The optimum size is 20 to 50, but the job of selecting such a small number for a truly « international » meeting is rather difficult; private or semi-private institutions have been more successful in organizing such small meetings than the Unions. However, many opportunities for small group meetings, formal and informal, are also provided during the big assemblies.

3. - Money raising and money distribution.

It is moderately easy to obtain money for science, provided one can make a good case backed by generally recognized scientific authority. Such authority exists in the Scientific Unions. For instance, much money flows from UNESCO to the International Scientific Unions, and from there the subsidies for travel to meetings, for publication and for scientific projects. True scientists have to be involved in this process, for only they have a chance of distinguishing between obsolete plans and fruitful ones. Much more could be said, but I must be brief.

4. - Terminology, units, codes.

Frequent contacts and reference to well-written papers or books go a long way towards establishing uniform terminology. Arabic numbers came to

western science by way of the latin translation of the *Almagest* in the Middle Ages. But long discussions are sometimes needed to reach agreement. Results range from minor matters, like the convention to reproduce spectra with the long wavelengths to the right, to major matters, like the general use of the metric system. (It is a shame to see still distances expressed in miles in scientific papers which are not meant for perusal by members of Congress). A recent example: a flow from west to east is called an east current in the ocean, a west wind in meteorology and an east wind in the ionosphere. But the dispute will be settled.

Another problem is that of telegraphic codes for messages about satellite orbits. In the few years since the first satellite was launched several independent code systems have come into use. The experts in the relevant COSPAR working group (Mrs. MASSEVITCH in the U.R.S.S., Dr. SHAPLEY in the U.S.A., and others) are well determined to reach uniformity. Again it may be clear that this is not a job which can easily be left in the hands of a committee of government officials.

5. - Joint projects.

Major joint projects need both the governmental and the non-governmental approach. A plan drawn up exclusively by scientists may be scientifically sound but completely impractical because commitments for major expenditures by the partners cannot be made. On the other hand, a purely inter-governmental plan may turn out to be impossible or hollow from the scientific point of view. It is advisable not to invite too many countries at once. The joint undertaking of countries of western Europe in C.E.R.N. has been quite successful. Inspiration from C.E.R.N. has been derived for the plans of E.S.O. and E.S.R.O., which are now on their way as temporary intergovernmental committees preparing a definite convention.

Another highly successful undertaking was the International Geophysical Year. After the end of the I.G.Y. it was decided to create some special committees in I.C.S.U. to carry on the task of certain disciplines in C.S.A.G.I. One of these is COSPAR. The problem of making direct national commitments possible (which was a weak point in the I.G.Y.) appears to have been solved successfully in COSPAR by adopting a somewhat complicated set of rules, in which the Union representatives have the major voice in determining scientific policy but the national representatives have exclusive rights to dispose of the money. World-wide cooperation, including the U.S.S.R., presents no major problems.

6. - International reference atmosphere.

About a year ago there was some correspondence about the possible need for a standard table of atmospheric pressures, densities and temperatures which would have international standing. The only internationally adopted « standard » was the I.C.A.O. atmosphere reaching to 22 km height; American Tables to 800 km were also widely used. COSPAR established a very small *ad hoc* group to study the problem, and via some ups and downs has reached the stage at which now, before the end of August, a set of Tables (of about 200 pages) will be published. This work has not gone through a process of « adoption », and therefore is not a « standard », but will be a very useful reference volume and probably represents about the best integrated Table which could be made with the data available all over the world as of early 1961.

7. - Frequency allocations.

I have left the most important points to the end. Frequency allocation is not just another example of how some agreement may make scientific work more convenient. It is vital in the literal sense: a matter of life or death. The science of radio astronomy, and equally well the purely scientific space research, would be dead before long if no frequency bands were kept free from interference for this work. Who can make and keep such bands free? The formal answer to this is: the national frequency committees in each country. Internationally, these committees adhere to agreements reached after extensive preparation in the I.T.U. and its organs such as C.C.I.R. and I.F.R.B. When the cities expanded and some observatories found that their work in optical astronomy had become impossible, the astronomers did not go to the town hall and ask: please dim the city lights. Instead, they found places far away from such interference. The situation in radio-astronomy is just as bad, but going away to quiet places helps very little (if we exclude the back side of the moon!). Hence, science must press its demands against the heavy weight of the commercial and military interests (among others) and in an ambience of a highly regulated intergovernmental organism. Here, again, the case cannot be won by sheer eloquence or clever administration: the scientists themselves must explain the situation, again and again, at different levels, must listen carefully to the counterarguments, and must show by detailed examples that their demands are justified and important.

At the Geneva conference in 1959, which lasted from mid-August to mid-December, and in which, for instance, the U.S.A. delegation counted 80 people,

these demands were pressed by a succession of scientists representing COSPAR, I.A.U. and U.R.S.I. In the Acts of this conference radio astronomy and space research have for the first time been recognized as services (which can receive allocations). For radio astronomy, only the 21 cm band obtained exclusive reservation. On many other bands the struggle must continue. And the worst is still to come: earth-to-earth communication via satellites for commercial, military, and other nonscientific purposes will claim large portions of the spectrum. The subject will come up during a big conference of the I.T.U., in 1963, for which arrangements are already being made. I.C.S.U. has therefore decided to create I.U.C.A.F. to maintain a continuous review of the situation, and to work out a good and timely strategy to keep radio astronomy and space science alive. At the same time, the developing countries may cease to be content with their relatively small share in a spectrum which is already overcrowded. Savings of frequencies can often be made, but only by doing away with obsolete equipment, which represents a financial as well as a psychological barrier. Clearly, a number of top scientists will have to give full attention to these problems, and I hope I have proven the point that they cannot be said to be wasting their time in doing so.

8. - Damaging experiments.

Experiments are now possible which may damage mankind or things held in esteem by mankind. Some such experiments are fully permitted, but others fully fatal. These problems cannot be answered by one simple formula, *e.g.* that nature should be preserved. For the continuation of human civilization requires certain artifacts.

Let us take a few examples of these dangerous experiments, briefly assembled in three categories.

Blasting. - A blast to make a road, or a mine, or an artificial earthquake for research into the structure of the earth's crust, is permissible. However, we should not blow up a city, or a continent, or the whole earth.

Seeding. - It has become accepted practice to blow puffs of sodium from a rocket to investigate the upper atmosphere. This may finally make it impossible to measure the natural sodium content. A much bigger seeding experiment was project « Argus », in August 1959, in which a few atom bombs temporarily created an artificial Van Allen belt. A plan for the near future is the project « West Ford » to create a belt of needles, which may hamper optical as well as radio astronomy. In all these cases we are approaching the borderline of the permissible.

Poisoning. - To poison the whole human population is possible but not permitted. To poison all life in Mars by germs brought from the earth is *perhaps* possible, even without our ever knowing what we had done. We feel this must be avoided, and extreme caution is needed in approaching the subject.

Each of the problems mentioned faces us with profound questions in the field of human relations as well as technically difficult scientific problems. I should like to urge the active scientists not to stay aloof from these problems through a mistaken professional pride, but to take full responsibility whenever an appeal is made for their insight. COSPAR as well as an earlier *ad hoc* committee, CETEX, have repeatedly emphasized this point.

Tipografia Compositori - Bologna - Italy
